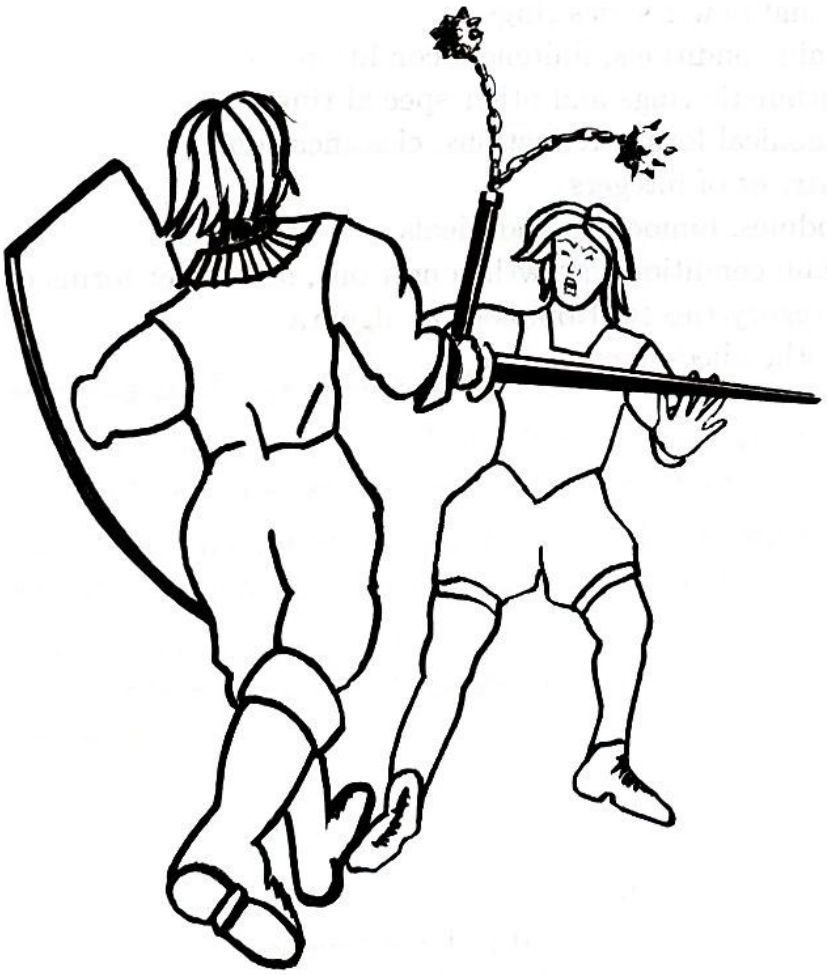
Grégory Berhuy

Algèbre : le grand combat

Troisième édition



Calvage & Mounet

Grégory Berhuy est professeur à l'université de Grenoble Alpes. Ses domaines de recherche de prédilection couvrent entre autres l'étude des invariants des structures algébriques, et l'application de l'algèbre non commutative à la communication sans fil.

gregory.berhuy@univ-grenoble-alpes.fr

Mathematics Subject Classification (2010) – Primary:

05XX Algebraic combinatorics
05E05 Symmetric functions and generalizations
11RXX Algebraic number theory: global fields
11R25 Quaternion and other division algebras: arithmetic, zeta functions
13AXX General commutative ring theory
13CXX Theory of modules and ideals
13F25 Formal power series rings
13EXX Chain conditions, finiteness conditions
13FXX Arithmetic rings and other special rings
15A21 Canonical forms, reductions, classification
15A36 Matrices of integers
16DXX Modules, bimodules and ideals
16PXX Chain conditions, growth conditions, and other forms of finiteness
18XX Category theory. Homological algebra
19AXX Grothendieck groups and K_0

Le dessin en page de titre est d'Isabella Bembo.
La « Greg's massue » est de Maya Tayara.

ISBN 978-2-493230-32-4

♾ Imprimé sur papier permanent

© Calvage & Mounet, Paris, 2025

Cet ouvrage est dédicacé :

— *à mes ~~victimes~~ étudiants passés, présents et futurs*

— *à mon ancien enseignant et collègue Georges Gras, qui m'a révélé la beauté de l'algèbre et l'existence des idéaux maximaux*

— *à mon ancien enseignant et collègue Henri Lombardi, qui m'a révélé la beauté de l'algèbre et la non-existence des idéaux maximaux*

— *à ma mère, fan de la première heure, et qui vante en la moindre occasion à qui veut bien l'entendre la qualité et l'utilité de mes ouvrages... comme presse-papiers.*

Table des matières

Partie 1. Phase d'approche : rappels et compléments

I. Théorie des ensembles

1. Applications 3
2. Relations d'équivalence 6
3. Lois internes, structures algébriques 11
4. Ensembles ordonnés 15
5. Exercices 16

II. Arithmétique dans $\mathbb{Z}$

1. Divisibilité dans $\mathbb{Z}$ 21
2. Division euclidienne, algorithme d'Euclide 23
3. Entiers premiers entre eux 26
4. Décomposition en facteurs premiers 27
5. Les pgcd et ppcm d'une famille quelconque d'entiers 29
6. Arithmétique modulaire 30
7. Exercices 33

III. Rappels et compléments d'algèbre linéaire

1. Généralités 37
2. Familles libres, génératrices, bases 40
3. Somme d'espaces vectoriels 44
4. Matrices représentatives, changement de base 46
5. Formes linéaires, droites et hyperplans 48
6. Exercices 53

IV. Déterminant
1. Introduction éclair au groupe symétrique 59
2. Formes multilinéaires alternées, déterminant 64
3. Déterminant d'un endomorphisme 67
4. Déterminant d'une matrice 69
5. Transvections et dilatations 75
6. Exercices 86

V. Un peu de géométrie
1. Affinités vectorielles 91
2. Espaces euclidiens, espaces hermitiens 95
3. Isométries d'un espace euclidien : premiers résultats 101
4. Structure des endomorphismes normaux 108
5. Espaces euclidiens orientés 114
6. Exercices 121

Partie 2. Groupes : il faut agir !!

VI. Propriétés élémentaires des groupes
1. Généralités 127
2. Sous-groupes 134
3. Sous-groupes engendrés par une partie 143
4. Théorème de Lagrange, ordre d'un élément 146
5. Interlude : formule d'inversion de Möbius 154
6. Groupes monogènes, groupes cycliques 156
7. Exercices 159

VII. Groupes opérant sur un ensemble
1. Actions de groupe 171
2. Premières applications 179
3. Coloriages 184
4. Exercices 193

VIII. Groupe symétrique, groupe alterné
1. Préliminaires 201
2. Décomposition en produit de cycles 204
3. Systèmes de générateurs 214
4. Signature, groupe alterné 215
5. Structure des groupes symétrique et alterné 219
6. Exercices 222

IX. Groupes quotients

1. Définition 237
2. Théorème de factorisation 241
3. Suites de composition 248
4. Présentations de groupes 264
5. Exercices 285

X. Produits directs et semi-directs

1. Préliminaires 297
2. Produits directs 299
3. Produits semi-directs 303
4. Exercices 313

XI. Actions de groupe et structure des groupes finis

1. Théorème de Sylow 317
2. Actions primitives et critère de simplicité d'Iwasawa 321
3. Exercices 334

XII. Thème et variations sur les groupes abéliens

1. Torsion dans un groupe abélien 347
2. Exposant d'un groupe 351
3. Caractères linéaires d'un groupe abélien fini 354
4. Groupes abéliens libres 358
5. Structure des groupes abéliens de type fini 364
6. Groupes abéliens divisibles 372
7. Exercices 384

XIII. Quand tout groupe fini d'ordre n est-il cyclique ?

1. Une condition nécessaire 389
2. Compléments de théorie des groupes 391
3. Le théorème principal 392

XIV. Petite classification des groupes finis ; épisodes précédents

1. Quelques théorèmes de classification 397
2. Groupes finis d'ordre $\leqslant 15$ 399

Partie 3. Anneaux : la loi des irréductibles

XV. Propriétés élémentaires des anneaux et des corps

1. Généralités 403
2. Éléments remarquables d'un anneau 410
3. Anneaux intègres, corps 414
4. Idéaux 419
5. Produit direct d'anneaux 426
6. Anneaux de polynômes 430
7. Polynômes symétriques 448
8. Exercices 459

XVI. Anneaux quotients, théorème chinois

1. Anneaux quotients, théorème de factorisation 471
2. Idéaux premiers, maximaux 475
3. Théorème chinois 479
4. Exercices 488

XVII. L'anneau $\mathbb{Z}/n\mathbb{Z}$ en long, en large et en travers

1. Échauffement 501
2. Structure de $(\mathbb{Z}/n\mathbb{Z})^\times$ 504
3. Carrés de $\mathbb{Z}/n\mathbb{Z}$ 508
4. Symboles de Legendre et de Jacobi 513
5. Exercices 527

XVIII. Divisibilité dans les anneaux

1. Divisibilité, éléments irréductibles et premiers 531
2. Arithmétique des anneaux principaux 540
3. Anneaux euclidiens 545
4. Anneaux noethériens 553
5. Anneaux factoriels 558
6. Exercices 566

XIX. Anneaux à division et codes Wifi

1. Le problème du codage 587
2. Recherche à corps perdu 591
3. Conservons notre énergie ! 596
4. Soyons discrets... 600

XX. Matrices à coefficients dans un anneau euclidien

1. Déterminant : le retour 603
2. Équivalence de matrices 606
3. Modules sur un anneau 618
4. Théorème de la base adaptée et applications 626
5. Exercices 638

XXI. Polynômes irréductibles et applications
1. Polynômes irréductibles et racines 643
2. Quelques critères d'irréductibilité 650
3. Factorialité de $A[X]$ 656
4. Irréductibilité d'un polynôme bicarré 663
5. Polynômes cyclotomiques et applications 674
6. Polynômes irréductibles sur $\mathbb{F}_p$, corps finis 684
7. Exercices 696

XXII. Idéaux de $A[X]$
1. Principalité des idéaux de $A[X]$ 701
2. Noethérianité de $A[X]$ 705
3. Idéaux premiers de $A[X]$, A principal 709
4. Exercices 713

XXIII. Séries formelles
1. Premières propriétés 715
2. Composition et dérivation de séries formelles 723
3. Séries formelles usuelles 739
4. Quelques applications des séries formelles 748
5. Exercices 754

XXIV. Complétion I-adique
1. Complété I-adique 761
2. Propriétés métriques 768
3. Complétion en un idéal principal 778
4. Exercices 788

Partie 4. Des corps ! Des corps partout !

XXV. Constructions à la règle et au compas et théorie des corps
1. Préliminaires 795
2. Extensions de corps 801
3. Extensions algébriques, transcendantes 812
4. Quelques problèmes de construction à la règle et au compas . . 821
5. Et après... 835
6. Exercices 837

XXVI. Polynômes et racines
1. Existence de racines dans une extension et applications 841
2. Corps de rupture, corps des racines d'un polynôme 850
3. Clôture algébrique d'un corps 857
4. Entiers algébriques 863
5. Exercices 869

XXVII. Plongements et extensions séparables
1. Plongements 878
2. Propriétés des extensions séparables 885
3. Polynômes séparables 891
4. Traces et normes 895
5. Exercices 904

XXVIII. Extensions galoisiennes
1. Extensions normales 911
2. Extensions galoisiennes 914
3. Le groupe de Galois d'un polynôme 921
4. Extensions cyclotomiques 929
5. Théorie de Kummer 934
6. Exercices 938

XXIX. Nul n'est censé ignorer Galois
1. Théorie de Galois 947
2. Résolubilité des équations par radicaux 956
3. Constructions à la règle et au compas : le retour 963
4. Exercices 965

Partie 5. Algèbre linéaire : réduisons en miettes !

XXX. Réduction des endomorphismes
1. Polynômes d'endomorphismes 975
2. Diagonalisation 988
3. Trigonalisation, décomposition de Dunford 995
4. Décomposition de Jordan 1010
5. Exponentielle de matrices 1021
6. Exercices 1032

XXXI. Décomposition de Frobenius

1. Sous-espace stable engendré par un vecteur 1047
2. Endomorphismes cycliques 1050
3. Décomposition de Frobenius : approche classique 1053
4. Décomposition de Frobenius : calcul pratique 1059
5. Commutant et bicommutant d'un endomorphisme 1072
6. Exercices 1077

XXXII. Dualité

1. Introduction 1083
2. Formes linéaires et bilinéaires, espace dual 1086
3. Interlude : espaces vectoriels quotient 1094
4. Dualité entre sous-espaces 1097
5. Bases duales 1104
6. Adjoint d'un endomorphisme 1109
7. Exercices 1117

XXXIII. Valeurs propres de matrices symétriques

1. Introduction 1125
2. La méthode générale 1129
3. Le cas $-1 \in k^{\times 2}$ 1132
4. Le cas $k = \mathbb{Q}$ 1137
5. Démonstration du théorème des quatre carrés 1143
6. Tout ce qui vous a été caché... 1145
7. Exercices 1151

Partie 6. Représentations : agissons dans l'espace !

XXXIV. Représentations linéaires des groupes finis : introduction

1. Modules simples, modules indécomposables 1157
2. Représentations linéaires : définitions et premiers exemples . . . 1161
3. Interlude : produit tensoriel 1176
4. Opérations sur les représentations 1181
5. Théorème de Maschke 1187
6. Exercices 1192

XXXV. Théorie des caractères

1. Caractère d'une représentation 1201
2. Orthogonalité des caractères et conséquences 1210
3. Table des caractères d'un groupe fini 1230
4. Décomposition en somme de représentations irréductibles 1237
5. Propriétés d'intégralité et applications 1246
6. Degré des représentations irréductibles 1254
7. Exercices 1261

XXXVI. Introduction au problème de Noether

1. Rappels et compléments sur les polynômes et les fractions rationnelles 1267
2. Polynômes et fractions rationnelles invariantes 1270
3. Le théorème de Fischer : aspect théorique 1280
4. Le théorème de Fischer : aspect pratique 1285

Bibliographie **1297**

Index **1301**

Avant-propos

Petit plaidoyer totalement partial pour l'étude de l'algèbre.– Nous voudrions commencer cet ouvrage par un petit plaidoyer en faveur de cette discipline mal-aimée des étudiants qu'est l'algèbre. Les étudiants se plaignent souvent de sa difficulté à cause de son haut niveau d'abstraction. Bien souvent, les objets manipulés ne leur paraissent pas assez concrets, et l'utilité de l'étude des structures algébriques, semblant absconses, leur échappe complètement. Pourquoi donc s'embêter à classifier les groupes finis d'ordre donné ou chercher toutes leurs représentations irréductibles, pourquoi s'acharner à savoir si tel groupe est résoluble ou si tel anneau est factoriel, pourquoi calculer la réduite de Jordan d'un endomorphisme et autres joyeusetés ? Pourquoi un tel besoin quasi irrépressible chez le mathématicien de généraliser à l'extrême, sinon pour torturer les étudiants par pur plaisir sadique [1] ?

Avant de tenter de répondre à ces questions, essayons de définir succinctement ce qu'est l'algèbre. La définition la plus juste, du point de vue où l'on se place, est sans doute la suivante : c'est l'étude des structures algébriques. Ici, on entend par structure algébrique un ensemble muni d'une ou plusieurs lois de composition internes ou externes, éventuellement muni d'un ordre ou une topologie, le tout satisfaisant un certain nombre d'axiomes. Ainsi, les groupes, les anneaux, les corps, les espaces vectoriels ou les représentations sont des structures algébriques. Une fois les structures définies, on définit souvent la notion de sous-structure, mais surtout les flèches entre deux telles structures, appelées « morphismes », qui sont « les applications compatibles avec les lois internes et externes ». Ce sont d'ailleurs les morphismes qui constituent la part importante de la théorie, plus que les objets eux-mêmes, puisque ce sont eux qui établissent les relations entre les objets, à travers la notion de structure quotient, par exemple. Une notion cruciale découlant directement de la notion de morphisme est la notion d'isomorphisme. Elle permet d'identifier des objets a priori différents mais structurellement identiques, et jouissant alors de propriétés similaires à tout

1. D'aucuns diront que la réponse est contenue dans la question.

point de vue. De manière un peu vague, deux structures isomorphes sont en fait deux façons de voir le même objet, mais dont on aurait étiqueté les éléments différemment. Par exemple, tous les espaces vectoriels de dimension n sur un même corps sont isomorphes, et une propriété vraie pour l'un est aussi vraie pour tous les autres, même si leurs éléments sont de natures différentes. D'ailleurs, après le choix d'une base, tout espace vectoriel de dimension n se comporte comme K^n : l'étiquetage des éléments se fait via les coordonnées dans la base choisie, mais le comportement des éléments vis-à-vis des opérations est exactement le même que celui des vecteurs de K^n muni de ses opérations usuelles. Un des problèmes fondamentaux en algèbre est alors de classifier les objets à isomorphismes près. Pour résumer, ce qui compte, ce n'est pas tant les objets, mais les flèches entre les objets [2].

Passons maintenant aux difficultés d'appréhender cette théorie. En ce qui concerne le problème du manque de palpabilité des objets, cela provient en grande partie d'un manque de pratique de l'abstraction, les étudiants étant confrontés aux objets abstraits et aux joies de la démonstration de théorèmes formels extrêmement tardivement. Autrement dit, pour maîtriser une nouvelle notion/discipline, il faut s'exercer. Ce principe frappé au coin du bon sens s'applique bien évidemment dans la vie de tous les jours (pratique d'un sport, d'un instrument, d'une langue, de la danse, etc.). Il n'y a aucune raison pour les mathématiques y échappent. Cela dit, la difficulté avec l'algèbre, et de manière plus générale avec les mathématiques, est que les définitions et théorèmes exposés en cours ou dans les ouvrages sont le fruit d'un long cheminement intellectuel, qui s'est déroulé parfois sur plusieurs dizaines d'années. Par conséquent, on ignore souvent le problème concret qui a motivé l'introduction de telle ou telle notion à l'origine. Par exemple, l'algèbre linéaire trouve essentiellement sa source dans l'étude des systèmes d'équations dites linéaires et également dans la géométrie. C'est Descartes qui fût le premier à utiliser des coordonnées pour résoudre des problèmes géométriques comme la détermination de l'intersection de deux droites, en termes d'équation linéaire, établissant dès lors un pont entre deux branches mathématiques jusqu'alors séparées : l'algèbre et la géométrie. Il faut attendre le dix-neuvième siècle et Gauss pour qu'une étude approfondie des systèmes linéaires soit faite. Camille Jordan résout également définitivement le problème de la réduction d'endomorphisme. Il faut noter que la notion d'espace vectoriel n'avait toujours pas été dégagée à ce moment là ! Une première tentative de formalisation de cette notion a été effectuée par Grassmann en 1844, dans son traité « La théorie de l'extension linéaire ».

2. Ce point de vue a d'ailleurs été poussé à l'extrême et avec succès à travers la théorie des catégories.

Ce traité, passé relativement inaperçu, contient déjà l'essentiel de la théorie telle que nous la pratiquons aujourd'hui. Enfin, en 1888, Peano, en se basant sur les travaux de Grassmann, axiomatise définitivement la théorie. On pourrait constater la même chose avec d'autres branches de l'algèbre. La théorie des groupes trouve ainsi son origine première dans le problème de résolubilité des équations polynomiales par radicaux et les travaux de Galois. La notion de groupe est également utilisée par Klein pour étudier les nouvelles géométries ayant émergé (géométrie hyperbolique, géométrie projective). Kummer introduit également un groupe abélien dans ses tentatives de démonstration du grand théorème de Fermat. Néanmoins, il faudra attendre 1882 pour qu'émerge la première définition axiomatique d'un groupe abstrait. Notons au passage que l'étude des anneaux prend également sa source dans les travaux de Dedekind sur le théorème de Fermat. C'est finalement Noether qui établira les fondements de la théorie des anneaux commutatifs unitaires. [3]

Vu la diversité des contextes dans lesquels toutes ces notions sont apparues historiquement, on comprend mieux pourquoi la mise en place d'une théorie générale est indispensable. D'ailleurs, établir une théorie unificatrice possède plusieurs avantages : éclairer sous un nouveau jour les cas particuliers qui ont motivé son élaboration, obtenir des résultats valables pour tous les objets de la théorie considérée (ce qui évite de redémontrer la même chose dans des dizaines de cas particuliers), mettre à disposition de nouveaux outils plus puissants pouvant être utiles pour démontrer de nouveaux théorèmes, créer des passerelles entre diverses branches des mathématiques, etc. Le prix à payer est souvent une montée dans l'abstraction, mais c'est un mal nécessaire pour pouvoir utiliser une artillerie efficace. Une image qui fonctionne plutôt bien est celle du randonneur complètement perdu : afin de retrouver plus facilement son chemin, il va chercher à prendre de la hauteur et à atteindre un point de vue plus élevé, pour avoir accès à un panorama complet des alentours, avant de redescendre pour revenir sur le bon chemin. Un des exemples les plus frappants est celui du grand théorème de Fermat : pour arriver finalement à une démonstration, il a fallu étudier les structures algébriques d'objets dont on aurait eu peine à imaginer l'existence au temps de Fermat. C'est ainsi que la théorie des anneaux, la théorie des nombres et de nouveaux pans de la géométrie algébrique se sont développés. Comme dans d'autres situations en mathématiques, le fait d'intégrer le problème de Fermat dans un cadre plus général et apparemment beaucoup plus difficile a permis de grandes avancées, parce que l'on dispose alors de tout un outillage développé pour ce cadre.

Essayons finalement de répondre à la question que beaucoup d'étudiants se posent légitimement : à quoi ça sert ? Mentionnons tout de suite la princi-

3. Toutes ces informations historiques sont tirées d'articles de Wikipédia.

pale force de l'algèbre, à savoir son ubiquité. Il est en effet intéressant de souligner que les diverses branches de l'algèbre ne sont pas compartimentées. Bien au contraire, il existe des passerelles entre elles et vers d'autres disciplines, ce qui permet une variété d'applications dans divers domaines.

La théorie des groupes illustre parfaitement notre propos. La notion de groupe apparaît naturellement dès qu'il s'agit d'étudier les symétries d'un objet. Elle intervient ainsi en géométrie puisque beaucoup d'exemples de groupes sont de nature géométrique, mais aussi en chimie (simplification des calculs de structure électronique, classification des états électroniques ou vibratoires d'une molécule ou d'un solide), en physique théorique (théorie de la relativité, théorie de l'unification) via la théorie des représentations. La théorie des groupes est aussi intimement liée à la théorie des corps via la théorie de Galois. Cette dernière théorie permet par exemple de donner une jolie caractérisation de la constructibilité d'un nombre à la règle et au compas, ou de la résolubilité d'une équation polynomiale par radicaux en termes de théorie des groupes. Une autre application de la théorie de Galois moins connue des étudiants (car plus pointue), mais omniprésente dans la recherche mathématique, est la résolution des problèmes dits « de descente ». De manière très floue, un problème de descente est un problème du type suivant : on considère un ensemble d'objets mathématiques et une relation d'équivalence sur ces objets. On se donne deux objets définis sur un corps K, et l'on suppose que les deux objets sont équivalents lorsque l'on les considère comme des objets définis sur un corps L contenant K. Les deux objets sont-ils équivalents sur K ? Par exemple, considérons le problème suivant : soient $M, M' \in \mathrm{M}_n(K)$ deux matrices. On suppose qu'il existe une matrice $Q \in \mathrm{GL}_n(L)$ telle que $M' = QMQ^{-1}$, où L est un corps contenant K. Existe-t-il une matrice $P \in \mathrm{GL}_n(K)$ telle que $M' = PMP^{-1}$? Il est bien connu que la réponse est positive, et la démonstration la plus classique ne nécessite pas de théorie de Galois (cf. corollaire XXXI-4.9). On peut alors légitimement se poser la question de savoir si c'est encore le cas lorsque l'on remplace le groupe linéaire par le groupe spécial linéaire ou le groupe orthogonal. La réponse est cette fois négative, et la théorie de Galois permet de caractériser précisément les matrices M' pour lesquelles cela devient vrai, à M fixé. Enfin, dans un registre plus ludique, la théorie des groupes permet de résoudre certains casse-têtes célèbres comme le Rubik's cube ou le jeu du taquin.

De manière plus générale, la théorie des groupes se révèle un outil indispensable dans l'étude des structures mathématiques, quelles qu'elles soient. En effet, pour comprendre un objet mathématique, il est toujours utile d'étudier son groupe d'automorphismes. Il est aussi naturel d'essayer de faire agir un groupe sur cet objet et d'étudier les orbites et stabilisateurs associés. On obtient alors à la fois de plus amples renseignements sur l'objet lui-même et

sur le groupe qui agit. Enfin, lorsque l'on s'attaque au problème de classification d'objets mathématiques à isomorphisme près, problème qui s'avère extrêmement ardu en général, il est toujours utile d'associer un groupe à chaque objet, de sorte que deux objets isomorphes aient des groupes associés isomorphes. Cette approche est souvent fructueuse, surtout lorsque le groupe associé est abélien, puisque les groupes abéliens sont des objets particulièrement bien compris et dans lesquels il est facile de calculer. Ce principe est illustré de manière extrêmement frappante en topologie. À tout espace topologique X et tout point $x \in X$, on peut associer un groupe abélien $H_k(X, X \setminus \{x\})$ pour tout entier $k \geqslant 0$, de sorte que, si $h : X \xrightarrow{\sim} Y$ est un homéomorphisme, on a un isomorphisme de groupes

$$h_* : H_k(X, X \setminus \{x\}) \xrightarrow{\sim} H_k(Y, Y \setminus \{h(x)\}).$$

Ces groupes peuvent être calculés dans certains cas, grâce à des théorèmes généraux. Par exemple, on peut démontrer que $H_k(\mathbb{R}^n, \mathbb{R}^n \setminus \{x\})$ est trivial si $k \neq n$, et est isomorphe à $\mathbb{Z}$ si $k = n$, et cela pour tout $x \in \mathbb{R}^n$. On en déduit aisément le résultat non trivial suivant, dont il serait difficilement concevable (voire impossible) d'avoir une démonstration à la main : pour tous entiers $n, m \geqslant 1$, les espaces topologiques $\mathbb{R}^n$ et $\mathbb{R}^m$ sont homéomorphes si, et seulement si, $n = m$.

L'algèbre linéaire quant à elle s'invite aussi bien dans les autres branches de l'algèbre qu'en analyse (équations différentielles ou aux dérivées partielles, espaces fonctionnels), en mécanique, en théorie des graphes, en théorie du codage (codes linéaires auto-correcteurs), dans l'étude des systèmes dynamiques discrets, en probabilités (chaînes de Markov), en informatique, etc. C'est aussi un excellent outil pour étudier les groupes, grâce à la théorie des représentations linéaires. Par exemple, la table des caractères permet de déterminer tous ses groupes distingués ou de décider de la simplicité d'un groupe (cf. théorèmes XXXV-1.8 et XXXV-1.10), ou de démontrer des théorèmes de structure (cf. théorèmes XXXV-5.4 et XXXV-5.5). Mentionnons aussi le théorème XXXV-2.20, dont la seule démonstration connue utilise la théorie des représentations linéaires. D'autre part, comme déjà mentionné plus haut, la théorie des représentations a des applications en chimie et en physique théorique.

Enfin, la théorie des anneaux possède un pouvoir d'ubiquité similaire à la théorie des groupes ou de l'algèbre linéaire. Cette théorie trouve son origine en théorie des nombres, avec les travaux de Dedekind visant à démontrer le théorème de Fermat, le point de départ étant que si x, y, $z \in \mathbb{Z}$ et p est un nombre premier impair, l'égalité $x^p + y^p = z^p$, peut se retraduire dans l'anneau $\mathbb{Z}[\zeta_p] = \{P(\zeta_p) \mid P \in \mathbb{Z}[X]\}$ sous la forme

$$(x + y)(x + \zeta_p y) \cdots (x + \zeta_p^{p-1} y) = z^p,$$

où $\zeta_p = e^{\frac{2i\pi}{p}}$.

La notion d'anneau intervient également de manière déterminante en géométrie algébrique, mais aussi en logique formelle (algèbres de Heyting, de Boole, algèbres cylindriques, etc.), en théorie des représentations linéaires (si V est une représentation irréductible d'un groupe G, l'anneau des endomorphismes de V est un anneau à division), en algèbre linéaire (théorie des polynômes annulateurs). Récemment, la théorie des anneaux à division a trouvé des applications en théorie des codes Wifi, et a servi à construire des codes performants pour la transmission de données sans fil (téléphone portable, box Wifi). Enfin, la théorie des anneaux a permis d'étudier globalement les formes quadratiques sur un corps K, en introduisant l'anneau de Witt $W(K)$, ce qui a insufflé un nouvel essor à la théorie des formes quadratiques et a contribué à la découverte de nombreux résultats.

De manière générale, l'utilisation des structures algébriques permet souvent de démontrer des résultats de manière simple et élégante, une fois que l'on a les bons outils (comme la résolution d'un problème de constructibilité à la règle et au compas, qui est purement géométrique, grâce à la théorie des corps et des groupes, ou comme la résolution d'une question topologique difficile à coup de groupes d'homologie). On pourrait continuer indéfiniment la liste des applications des notions présentées dans cet ouvrage. Nous espérons que ce panorama réduit suffira à convaincre le lecteur du bien-fondé de l'étude des structures algébriques et de la beauté de cette théorie.

Prérequis.–

Le texte qui suit est un cours d'algèbre s'adressant à des lecteurs ayant acquis les notions mathématiques de base de première et deuxième années de Licence. Même si des rappels seront faits dans la première partie, nous supposerons que le lecteur maîtrise quelques notions simples de théorie des ensembles, l'arithmétique dans $\mathbb{Z}$, et surtout l'algèbre linéaire et matricielle de base. Le lecteur ne se sentant pas au point sur tout ou partie de ces notions est vivement encouragé à d'abord consolider ses connaissances sur les points précédents avant de s'attaquer à la lecture de cet ouvrage.

À qui s'adresse ce livre ?–

Ce livre est une version extrêmement étoffée de notes de cours d'algèbre donnés par l'auteur en L3 et en M1 à l'université de Grenoble Alpes. Il couvre entre autres l'intégralité du bagage algébrique minimum que devrait avoir acquis un étudiant en L3 et en M1 (sans parler des agrégatifs !). Il contient également des résultats qui permettent d'approfondir certaines thématiques ou de redécouvrir des notions classiques sous une nouvelle perspective, et qui pourront intéresser les candidats à l'agrégation interne ou

externe. Aussi, certains chapitres s'adressent-ils spécifiquement aux étudiants de M1, d'autres encore aux agrégatifs. La répartition, totalement subjective, du contenu de cet ouvrage par niveaux peut se résumer comme suit.

• **L3**

- toute la partie 1 ;

- l'intégralité des chapitres VI, VII et VIII, exception faite de la section §3, qui peut être passée sans que cela ne nuise à la compréhension de la suite ;

- les sections §1 et §2 du chapitre IX ;

- le chapitre X, au moins jusqu'au corollaire X-3.11 ;

- la section §1 du chapitre XI ;

- la section §2 du chapitre XII, et éventuellement, à titre de compléments, la section §3 et le théorème XII-5.3 ;

- le chapitre XIV ;

- le chapitre XV, au moins jusqu'à l'exemple XV-6.32 inclus ;

- le chapitre XVI dans son intégralité, sauf peut-être le lemme XVI-2.4 et le théorème XVI-2.5, qui peuvent être passés en première lecture ;

- les sections §1 et §2 du chapitre XVII ;

- les sections §1, §2 et §3 du chapitre XVIII ;

- les sections §1, §2, §3 et §4 du chapitre XXX, et le début de la section §5 jusqu'au lemme XXX-5.6 inclus ;

- pour la culture, les chapitres XIII , XIX et XXXIII.

• **M1**

- tout ce qui précède ;

- la section IX-3, du moins ce qui concerne les groupes résolubles (le théorème de Jordan-Hölder pouvant être passé) ;

- l'intégralité des chapitres XV, XVI,XVIII, XX ;

- le chapitre XXI, sauf la section §4, qui est extrêmement technique et dont l'importance des résultats est plutôt anecdotique ;

- éventuellement le chapitre XXIII ;

- l'intégralité des chapitres XXV à XXIX ;

- les sections §1, §2 et §4 du chapitre XXXI ;

- l'intégralité des chapitres XXXII à XXXV ;

- pour la culture, le chapitre XXXVI.

• **Agrégation**

- l'intégralité de ce livre.

Notons néanmoins que cet ouvrage n'est **pas** un livre de préparation à l'agrégation. Tous les résultats s'adressant spécifiquement aux agrégatifs dans cet ouvrage (i.e. ceux contenus dans les chapitres/sections non mentionnés ci-dessus) doivent plutôt être vus comme des prolongements naturels aux notions abordées, et destinés à étoffer la culture générale mathématique du lecteur, même si certains peuvent être utilisés comme des points de développement dans les leçons (comme ceux concernant le symbole de Legendre-Jacobi, la démonstration classique de l'existence de la décomposition de Frobenius, ceux exposés dans le chapitre XXII, ou même ceux de certains exercices). D'autres résultats en revanche, comme ceux sur l'irréductibilité de certains polynômes bicarrés (Chapitre XXI, §4), ont probablement des démonstrations trop longues et/ou trop techniques pour pouvoir être utilisés comme points de développement, mais méritaient néanmoins d'avoir leur place dans cet ouvrage, pour leur originalité ou leur intérêt mathématique.

Contenu et notes bibliographiques.–

La première partie propose des rappels et de compléments. Dans le chapitre I, on fait de brefs rappels sur les applications, ainsi que sur les relations d'équivalence et les relations d'ordre. On introduit aussi brièvement les structures algébriques (groupes, anneaux, corps) qui seront étudiées dans cet ouvrage. Le chapitre II fournit des rappels de base sur l'arithmétique dans $\mathbb{Z}$ et l'arithmétique modulaire. Au passage, on introduit l'anneau $\mathbb{Z}/n\mathbb{Z}$, qui sera étudié plus en profondeur dans la troisième partie. On pourra retrouver toutes les notions abordées dans les chapitres I et II dans [3]. Le lecteur désireux d'en savoir un peu plus sur la théorie des ensembles pourra se référer à [33] ou [42]. [3]. En particulier, on trouvera une construction axiomatique de $\mathbb{N}$ dans les références précédentes. Dans le chapitre III, on rappelle les résultats d'algèbre linéaire qui seront utiles dans toute la suite, en proposant des généralisations au cas des espaces vectoriels de dimension infinie lorsque c'est possible. On étudie également succinctement les formes linéaires. Dans le chapitre IV, après une introduction éclair au groupe symétrique, on construit le déterminant d'un endomorphisme, et l'on démontre ses principales propriétés. On achève ce chapitre en étudiant les transvections et les dilatations d'un espace vectoriel de dimension finie. Le lecteur ayant besoin de se rafraîchir la mémoire sur les notions abordées dans ces chapitres pourra consulter [3] ou [31] avec profit. Le chapitre V expose quelques résultats classiques de géométrie vectorielle. Après avoir caractérisé les affinités vectorielles, on propose de brefs rappels sur

les espaces euclidiens et hermitiens. On détermine alors la structure des endomorphismes normaux d'un espace euclidien ou hermitien. En particulier, on obtient la structure des isométries de tels espaces, ainsi que le théorème de diagonalisation d'une matrice symétrique réelle en base orthonormée. Ce chapitre s'achève par quelques considérations sur les espaces euclidiens orientés. On définit la notion d'orientation, de base orthonormée, puis le produit mixte et le produit vectoriel. Le lecteur pourra retrouver tous les résultats sur les espaces euclidiens et leurs démonstrations dans [31].

La deuxième partie concerne la théorie des groupes. Le chapitre VI expose les notions élémentaires, à savoir groupes, sous-groupes, morphismes, sous-groupes engendrés par une partie, groupes cycliques et monogènes, ordre d'un élément et d'un sous-groupe. Le chapitre VII introduit une notion centrale dans la théorie, à savoir la notion de groupe agissant sur un ensemble, et qui fait le lien entre la théorie des groupes et la géométrie. On donne alors les premières applications des actions de groupe. La dernière section applique les résultats de ce chapitre à un problème de combinatoire. Dans le chapitre VIII, on étudie la structure du groupe symétrique et du groupe alterné (générateurs, sous-groupes distingués). Le chapitre IX introduit la théorie des groupes-quotient. Après avoir démontré le théorème de factorisation et ses corollaires, on s'intéresse ensuite aux diverses suites de composition associées à un sous-groupe. On démontre entre autres l'existence d'une suite de Jordan-Hölder, ainsi que quelques résultats classiques sur les groupes (super)-résolubles. Enfin, on introduit la notion de groupe libre et présentation d'un groupe par générateurs et relations. Dans le chapitre X, on introduit la notion de produit direct externe et interne, ainsi que la notion plus délicate de produit semi-direct. On démontre alors quelques critères simples pour déterminer si un groupe est isomorphe à un produit (semi-)direct. Le chapitre XI donne deux belles applications des actions de groupe à la structure des groupes finis : un critère de simplicité d'un groupe, dû à Iwasawa, et le théorème de Sylow, qui démontre l'existence de p-sous-groupes d'un groupe fini d'ordre maximal (où p est un nombre premier). Ce dernier théorème, joint aux résultats du chapitre précédent, permet déjà d'obtenir pas mal de résultats de classification de groupes finis d'ordre donné. Le chapitre XII s'articule autour des groupes abéliens, et contient des résultats classiques autour de cette notion : existence d'un élément d'un groupe abélien fini dont l'ordre est égale à l'exposant du groupe, structure des groupes abéliens de type fini, structure des groupes abéliens divisibles. Enfin, dans le chapitre XIII, on donne une caractérisation des entiers $n \geqslant 1$ tels que tout groupe d'ordre n est cyclique. Les notions et résultats abordés dans cette partie sont extrêmement classiques, et peuvent être retrouvés dans n'importe quel ouvrage consacré à la théorie des groupes, comme par exemple [24], [39] ou [49].

La troisième partie est consacrée à la théorie des anneaux. Dans le chapitre XV, on introduit la notion d'anneau, ainsi que ses éléments remarquables (diviseurs de zéro, éléments inversibles, nilpotents, idempotents). On donne aussi une condition nécessaire et suffisante pour qu'un anneau soit isomorphe au produit direct de deux anneaux non triviaux. On définit également l'anneau des polynômes à coefficients dans un anneau, et l'on étudie ses premières propriétés. Enfin, on démontre le théorème de structure des polynômes symétriques en n indéterminées. Le chapitre XVI s'intéresse aux anneaux-quotient, et aux résultats classiques inhérents à cette notion : théorème de factorisation, caractérisation des idéaux premiers et maximaux en termes d'anneau-quotient, théorème chinois. Dans le chapitre XVII, on étudie l'anneau $\mathbb{Z}/n\mathbb{Z}$ de manière approfondie. Après avoir déterminé les éléments nilpotents et idempotents de cet anneau, on décrit explicitement la structure du groupe des éléments inversibles. On s'intéresse ensuite aux carrés de $\mathbb{Z}/n\mathbb{Z}$. Au passage, on introduit les symboles de Legendre et de Jacobi, et l'on propose une démonstration de la loi de réciprocité quadratique fondée sur le calcul modulaire. Le chapitre XVIII est consacré à l'étude des propriétés arithmétiques des anneaux. On s'intéresse d'une part à la généralisation des notions de pgcd et ppcm, et d'autre part à la généralisation du théorème fondamental de l'arithmétique. Pour cela, on introduit la notion d'élément irréductible d'un anneau, et l'on étudie l'existence et l'unicité d'une factorisation d'un élément non nul en produit d'un inversible et d'éléments irréductibles, d'abord dans le cadre agréable des anneaux principaux, puis dans un cadre général. Cela conduit à la notion d'anneau factoriel. On démontre alors des analogues des résultats classiques de l'arithmétique. Au passage, on étudie brièvement les anneaux noethériens. Dans ce chapitre, on étudie aussi les anneaux euclidiens, c'est-à-dire les anneaux possédant un analogue de la division euclidienne. On généralise alors l'algorithme d'Euclide à de tels anneaux. Comme illustration des notions introduites dans les chapitres précédent, le chapitre XIX donne une application de la théorie des anneaux à division à la communication sans fil. Le chapitre XX est consacré à l'étude de l'équivalence de matrices à coefficients dans un anneau A (deux matrices de même taille étant équivalentes si on peut passer de l'une à l'autre par des opérations élémentaires sur les lignes et les colonnes). Lorsque A est euclidien, on montre que l'on peut réduire toute matrice à l'aide d'opérations élémentaires en une matrice d'une forme particulière, appelée forme normale de Smith. Après avoir brièvement introduit la notion de A-module (qui généralise de façon naturelle la notion d'espace vectoriel sur un corps), on en déduit alors une méthode systématique pour résoudre les systèmes linéaires à coefficients dans A. On montre aussi que, lorsque A est euclidien, le groupe $\mathrm{GL}_n(A)$ des éléments inversibles de l'anneau $\mathrm{M}_n(A)$ est engendré par les matrices de transvection et de dilatation. On applique aussi la théorie développée dans ce chapitre pour

démontrer de manière plus élégante le théorème de structure des groupes abéliens de type fini. Nous avons choisi sciemment de passer sous silence le théorème de structure des modules de type fini sur un anneau euclidien (ou même sur un anneau principal), y compris lorsque nous parlerons de décomposition de Frobenius dans le chapitre XXXI. Le lecteur intéressé par une exposition plus complète de la théorie des modules pourra se référer à [9], ainsi qu'aux exercices. Le chapitre XXI se concentre tout particulièrement sur les éléments irréductibles de l'anneau $A[X]$ des polynômes à coefficients dans un anneau A. On commence par souligner que cette notion est bien plus délicate que dans le cas des polynômes à coefficients dans un corps. On donne ensuite quelques critères d'irréductibilité classiques (critère de réduction modulo un idéal, critère d'Eisenstein). On montre ensuite que si A est factoriel, il en est de même de l'anneau $A[X]$. On montre ensuite les limites des critères d'irréductibilité démontrés précédemment, en exhibant une famille de polynômes de degré 4 dans $\mathbb{Z}[X]$ et $\mathbb{Q}[X]$, mais qui sont réductibles modulo tout idéal de $\mathbb{Z}$. On définit ensuite une famille très importante de polynômes irréductibles de $\mathbb{Q}[X]$, à savoir les polynômes cyclotomiques. On les utilise ensuite pour démontrer une version faible du théorème de la progression arithmétique de Dirichlet, et pour démontrer que tout anneau à division fini est commutatif. Enfin, on démontre l'existence de polynômes irréductibles de tous degrés sur $\mathbb{F}_p$, et l'on élucide la structure des corps finis. On décrit enfin l'algorithme de Berlekamp, qui permet de factoriser de manière redoutablement efficace les polynômes à coefficients dans $\mathbb{F}_p$ en produit de polynômes irréductibles. Dans le chapitre XXII, on s'intéresse plus particulièrement aux idéaux de l'anneau $A[X]$. On détermine en particulier sous quelles conditions ces idéaux sont de type fini/principaux. On décrit également tous les idéaux premiers et maximaux de $A[X]$ lorsque A est un anneau principal. Dans le chapitre XXIII, on étudie les propriétés des séries formelles à coefficients dans un anneau A. En particulier, on définit diverses opérations sur les séries formelles comme la dérivation ou l'évaluation. On introduit ensuite quelques séries formelles usuelles et l'on donne quelques applications des séries formelles à la combinatoire. Le lecteur pourra trouver d'autres applications des séries formelles dans [3] ou [1], ainsi que dans les exercices. Enfin, dans le chapitre XXIV, on définit de manière algébrique le complété d'un anneau A par rapport à un idéal I. Cette construction permet à la fois d'apporter un nouvel éclairage sur les propriétés des séries formelles, et de construire l'anneau des entiers p-adiques, un anneau extrêmement important en théorie des nombres. Au passage, on fait le lien avec la complétion d'un anneau par rapport à une distance, bien connue en topologie. La plupart des résultats traités dans cette partie peuvent se retrouver dans [1], [30] et [37]. Le lecteur désirant approfondir ses connaissances sur la théorie des nombres p-adiques pourra consulter [4] ou [29].

La quatrième partie est consacrée à la célèbre théorie de Galois. Dans le chapitre XXV, on commence par motiver l'étude des extensions de corps en montrant comment ces objets interviennent naturellement dans un problème a priori purement géométrique, à savoir le problème de la construction à la règle et au compas. Ces considérations nous conduisent à introduire la notion d'élément algébrique sur un corps. Après avoir montré quelques résultats de théorie des extensions, nous résolvons des problèmes classiques de construction à la règle et au compas, en particulier la construction des polygones réguliers. Nous achevons ce chapitre en expliquant comment ces constructions amènent tout naturellement à de nouvelles questions de théorie des corps, et à la nécessité de disposer d'un corps dans laquelle tout polynôme non constant de $K[X]$ se factorise en produit de polynômes de degré 1. Le chapitre XXVI rentre dans le vif du sujet, et se préoccupe de l'existence éventuelle de racines d'un polynôme non constant $P \in K[X]$ dans un corps L contenant K. Cela aboutit à la notion de corps de rupture et de corps des racines. On en déduit entre autres une nouvelle façon de démontrer l'existence de corps finis de cardinal donné. On montre également que $\mathbb{C}$ est algébriquement clos. On s'intéresse ensuite à l'existence et l'unicité d'une clôture algébrique d'un corps K. Enfin, on introduit la notion d'élément entier sur un anneau commutatif, ce qui nous permet ensuite de comparer les relations entre l'irréductibilité d'un polynôme dans l'anneau $A[X]$ et son irréductibilité dans l'anneau $K_A[X]$, où A est un anneau intègre et K_A est son corps des fractions. Le chapitre XXVII étudie les extensions de corps dites séparables. On montre alors que ces extensions ont de bonnes propriétés vis-à-vis du prolongement des morphismes de corps à valeurs dans un corps algébriquement clos. Dans le chapitre XXVIII, on introduit la notion d'extension galoisienne, et l'on donne quelques familles classiques d'extensions galoisiennes (extensions cyclotomiques, extensions de Kummer). Dans le chapitre XXIX, on établit une correspondance bijective entre l'ensemble des sous-extensions d'une extension galoisienne et l'ensemble des sous-groupes de son groupe d'automorphismes. On revient alors sur le problème de construction à la règle et au compas. Pour finir, on s'intéresse au problème qui est historiquement à l'origine de la théorie de Galois et de la théorie des groupes, à savoir le problème de la résolubilité d'une équation polynomiale par radicaux. Le lecteur désireux de découvrir d'autres ouvrages consacrés à la théorie de Galois pourra consulter [18], [18], [46] ou [47], ainsi que [48] pour un aspect historique complet de la question de la résolubilité des équations par radicaux et des travaux d'Évariste Galois.

La cinquième partie de cet ouvrage est essentiellement dévolue à la réduction des endomorphismes. Dans le chapitre XXX, après avoir introduit la notion de polynôme annulateur, on donne les critères classiques de diagonalisabilité et trigonalisabilité d'un endomorphisme, puis on s'intéresse à la

décomposition de Dunford et à la réduction de Jordan. La dernière partie du chapitre est consacrée à l'étude de l'exponentielle de matrices. Le chapitre XXXI établit l'existence d'une décomposition d'un espace vectoriel V en somme directe de sous-espaces u-cycliques (où u est un endomorphisme de V). Cela conduit à la notion d'invariants de similitudes d'un endomorphisme. L'existence d'une telle décomposition est établie de deux manières, d'une part en utilisant la dualité, et d'autre part en utilisant l'algorithme de réduction des matrices à coefficients dans l'anneau euclidien $K[X]$. On donne alors un algorithme qui permet de décider si deux endomorphismes de V donnés sont semblables ou non. On donne ensuite quelques applications de la décomposition de Frobenius à l'étude du commutant et du bicommutant d'un endomorphisme. Enfin, dans le chapitre XXXII, on étudie la dualité en dimension quelconque. Plus précisément, étant donné une application bilinéaire $b : E \times F \longrightarrow K$, on donne des conditions nécessaires et suffisantes pour l'on ait une correspondance bijective entre les sous-espaces de E et ceux de F (via les sous-espaces orthogonaux à gauche et à droite relatifs à b). On étudie également l'existence de bases duales et d'endomorphismes adjoints. Le chapitre XXXIII fournit une jolie application de la dualité et de la théorie des extensions de corps à la caractérisation des valeurs propres de matrices symétriques. Les résultats sur la réduction des endomorphismes sont de grands classiques, et peuvent être retrouvés un peu partout dans la littérature, comme par exemple [1] ou [7]. Le lecteur retrouvera certains résultats du chapitre sur la dualité dans [5].

Enfin, la sixième et dernière partie de ce livre fournit une introduction à la théorie des représentations linéaires des groupes finis. Le chapitre XXXIV introduit la notion de représentation linéaire d'un groupe fini sur un corps, ainsi que la notion importante de représentation irréductible. On définit ensuite plusieurs opérations sur les représentations linéaires (somme, produit tensoriel, carrés symétrique et alterné, etc.). On montre alors que, sous certaines hypothèses, toute représentation linéaire se décompose en somme directe de sous-représentations irréductibles. L'ultime chapitre XXXV expose la théorie des caractères, qui est le cœur de la théorie des représentations. On démontre, entre autres, que lorsque le corps de base est algébriquement clos de caractéristique nulle, le caractère d'une représentation la caractérise à isomorphisme près. On démontre également (sous les mêmes hypothèses) certaines relations d'orthogonalité entre les caractères des représentations irréductibles. On donne aussi quelques applications de la théorie des caractères à la théorie des groupes, dont le théorème $p^a q^b$ de Burnside. On fournit enfin quelques résultats sur le degré des représentations irréductibles, ainsi qu'une méthode explicite pour décomposer une représentation donnée en somme directe de sous-représentations irréductibles. Le chapitre XXXVI donne une brève introduction au problème de

Noether, qui concerne la structure des fractions rationnelles invariantes sous l'action d'un groupe fini. Nous n'avons pas abordé la notion de représentation induite, notion très importante pour construire des représentations irréductibles via le critère de Mackey. Le lecteur désireux d'approfondir ce sujet pourra consulter [2], [20], [35], [38] ou [45]. Un autre thème intéressant est la détermination des représentations irréductibles du groupe symétrique. Le lecteur curieux pourra se référer à [2], [35] ou [38] pour de plus amples renseignements. Tous les résultats de cette partie peuvent se retrouver dans les références mentionnées ci-dessus.

Notons également que la majeure partie du contenu de cet ouvrage peut être retrouvée dans [6], [20], [30], [34] ou [37].

Chaque chapitre (ou presque) s'achève par une série d'exercices de niveaux variés. En ce qui concerne les chapitres mélangeant plusieurs niveaux, les exercices nécessitant une connaissance des sections spécifiques au niveau M1 ou Agreg[4] seront signalés. Nous encourageons le lecteur à essayer de les résoudre, quitte à sécher dessus un bon moment en ce qui concerne les exercices un peu plus difficiles, car comme disait Paul Langevin : « le concret, c'est de l'abstrait rendu familier par l'usage ».

Quelques conseils pour travailler son cours.–

Les conseils qui suivent sont évidemment valables pour n'importe quel cours, qu'ils soit suivi en amphi ou qu'il soit étudié en autodidacte à l'aide d'un livre. Chaque étudiant est bien entendu différent, et il n'y a pas de recette universelle pour maîtriser un cours de mathématiques. Néanmoins, il y a quand même quelques conseils frappés au coin du bon sens[5] qui peuvent faciliter l'apprentissage.

(1) S'aménager des conditions de travail propices (que ce soit seul ou en groupe) : couper télé, radio, téléphone portable, internet, afin de pouvoir mieux se concentrer.

(2) Apprendre les définitions et les théorèmes principaux à la virgule près (chaque mot et l'ordre de chaque mot ayant son importance dans un énoncé mathématique). Connaître les exemples classiques des objets que l'on a défini.

(3) Savoir faire les exercices d'application directe du cours.

(4) S'attaquer aux exercices moins évidents, quitte à sécher dessus un bon moment. On apprend les maths en les faisant (et l'on éprouve généralement une grande satisfaction personnelle lorsque l'on trouve finalement). Il est absolument inefficace d'abandonner au bout de dix minutes. Pour progresser, il faut être patient et tenace.

4. Suivant le découpage présenté plus haut.
5. Tout du moins selon l'auteur...

(5) Une fois cela fait, et seulement à ce moment-là, on peut reprendre chaque énoncé important du cours et essayer de répondre (dans la mesure du possible) aux questions suivantes : pourquoi a-t-on fait telle ou telle hypothèse ? Où intervient-elle dans la démonstration ? Le résultat est-il encore vrai si l'on affaiblit les hypothèses, ou si l'on en enlève ? Si ce n'est pas le cas, peut-on produire un contre-exemple ?
C'est très bien pour comprendre les choses en profondeur.

(6) Si possible, essayer de comprendre les étapes essentielles et l'articulation d'une démonstration d'un théorème important. Savoir refaire une démonstration courte.
Il faut prendre conscience que l'on ne peut pas tout comprendre à fond sur le moment, et accepter que l'on puisse ne pas maîtriser une démonstration tout de suite. Il faut parfois un an de recul sur une notion pour vraiment la digérer.

(7) Souvent, quand on lit une démonstration du cours ou la solution d'un exercice, on a le sentiment d'avoir bien compris. Il faut savoir que neuf fois sur dix, il n'en est rien. Pour voir si c'est vraiment le cas, nous conseillons au lecteur d'essayer de ré-expliquer la dite démonstration ou solution à des collègues. C'est un test qui ne trompe pas : s'il y avait des points obscurs ou doutes, cela se ressentira lors de la tentative d'explication. Comme on dit souvent : ce qui se conçoit bien s'énonce clairement.

Comment aborder un exercice.–

Enfonçons le clou : on apprend les maths en les faisant, et en se trompant, et certainement pas en lisant linéairement un ouvrage, une page Web ou une solution rédigée d'exercice. Voici quelques principes de base pour éviter de rester bloqué devant un exercice.

(1) Lire l'énoncé en entier pour identifier le but de l'exercice[6].

(2) Identifier de quoi ça parle, et lister les théorèmes à disposition sur le sujet.

(3) Repérer les hypothèses, ainsi que les mots du style « en déduire ». Cela donne une bonne indication sur la manière de procéder.

(4) Lorsque que l'on doit faire une démonstration, ne pas hésiter à écrire les hypothèses, et la conclusion, pour visualiser d'où l'on part et où l'on veut arriver. Cela aide souvent pour savoir comment démarrer.

(5) Lorsque l'on rédige une démonstration, il faut savoir justifier chaque étape. Si ce n'est pas le cas, c'est qu'il y a un problème...

6. Dit comme ça, le conseil a l'air complètement idiot, mais c'est en fait vital pour ne pas perdre de vue le but final !

(6) Bien faire apparaître où l'on utilise les hypothèses de l'énoncé dans la rédaction d'une démonstration. Si une hypothèse n'est pas utilisée, il y a deux possibilités : soit l'énoncé est mal fait (ce qui peut arriver), soit il y a une ou plusieurs erreurs dans la démonstration en question (ce qui est le cas le plus probable).

(7) Se forcer à bien rédiger et à écrire des phrases mathématiquement complètes et grammaticalement correctes, **même au brouillon.** Il est extrêmement facile de prendre de mauvaises habitudes de rédaction.

Parmi les principes de bonne rédaction, on peut citer : ne pas utiliser d'abréviation, ne pas mélanger symboles mathématiques et français, n'écrire un signe « $\Longleftrightarrow$ » que si l'on est certain d'avoir une équivalence logique, dire où l'on utilise les hypothèses, vérifier que l'on n'utilise pas la conclusion souhaitée pour la démontrer, etc.

(8) Faire des exercices dont on ne dispose pas du corrigé, pour ne pas être tenté d'aller y jeter un coup d'œil [7].

(9) Comme déjà évoqué précédemment, essayer d'expliquer sa proposition de solution d'un exercice à des collègues permet de déceler des erreurs ou des points mal rédigés.

Remerciements.–

L'idée d'écrire cet ouvrage m'est venue lorsque j'ai eu vent à plusieurs reprises que certains de mes anciens étudiants utilisaient encore mes polycopiés de cours de L3 et de M1 pour préparer l'agrégation de mathématiques, mais qu'ils n'étaient hélas pas autorisés à utiliser les dits polycopiés pour passer leurs épreuves orales. Je les remercie donc pour m'avoir encouragé directement ou indirectement à entreprendre cette trèèèèès longue entreprise de rédaction. Merci aussi à toutes mes victimes estudiantines, dont les commentaires m'ont permis d'améliorer mes cours au fil des années. Je tiens aussi à remercier tous mes relecteurs. Sans leurs conseils avisés et leurs suggestions pertinentes, cet ouvrage ne serait pas ce qu'il est [8]. Merci donc à Vincent Beck, Alain Debreil, Christophe Derras, Sébastien Gerekens, Éloi Mehr, Éric Pite, Étienne Rousée et Rémi Saint-Romain.

Enfin, merci du fond du cœur à Alain Debreil, Rached Mneimné et à toute l'équipe de C&M pour le travail éditorial titanesque qu'ils ont accompli avec courage et abnégation, ainsi que pour leur patience d'ange face à mes caprices d'auteur.

7. Cela tombe bien, dans cet ouvrage, aucune correction n'est donnée : cela ne sert à rien, et ce n'est pas rentable.

8. En particulier, il n'y aurait pas eu de chapitres sur la complétion I-adique, ou sur la théorie de Galois.

La deuxième édition a été l'occasion de quelques ajouts et autres corrections, et d'une amélioration sensible de la mise en page.

Cette troisième édition a été l'occasion de corriger quelques coquilles, de modifier certaines démonstrations, d'ajouter de nouveaux exercices, et d'inclure quatre nouveaux chapitres permettant d'illustrer les notions introduites dans ce livre, à savoir les chapitres XIII, XIX, XXXIII et XXXVI.

Première partie

Phase d'approche : rappels et compléments

Chapitre I

Théorie des ensembles

Dans ce chapitre, on regroupe quelques définitions et résultats sur les relations d'équivalence et relations d'ordre qui nous seront utiles par la suite. On introduit également la notion de loi interne sur un ensemble, et l'on définit les groupes, anneaux et corps, qui nous serons utiles pour aborder sereinement la suite.

On commence par rappeler quelques notions relatives aux applications.

1. Applications

Commençons par donner la définition d'une correspondance entre deux ensembles.

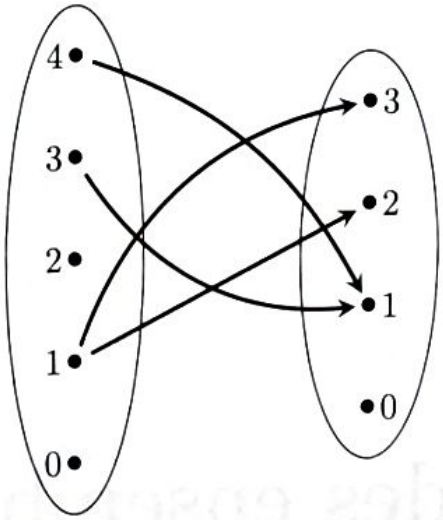
1.1. Définition. Une *correspondance* est un triplet $f = (E, F, R)$, où E et F sont deux ensembles, et R est une partie de $E \times F$. L'ensemble E est appelé *ensemble de départ (ou ensemble source)*, et F est appelé *ensemble d'arrivée (ou ensemble but).*

Si $x \in E$, *une image* de x par f est un élément $y \in F$ tel que $(x, y) \in R$.

Si $y \in F$, *un antécédent* de y par f est un élément $x \in E$ tel que $(x, y) \in R$.

Un bon moyen pour comprendre cette définition est de se représenter une correspondance sous forme de diagramme, à l'aide de patates et de flèches (au moins lorsque E et F sont finis).

1.2. Exemple. Un tel exemple de correspondance est donné en page suivante entre l'ensemble $E = \{0, 1, 2, 3, 4\}$ et l'ensemble $F = \{0, 1, 2, 3\}$.



Ici, on a $R = \{(1,2),(1,3),(3,1),(4,1)\}$. Les éléments $0, 2 \in E$ n'ont pas d'image dans F, tandis que $1 \in E$ a deux images dans F. De plus, $0 \in F$ n'a pas d'antécédent dans E, tandis que $1 \in F$ a deux antécédents dans E.

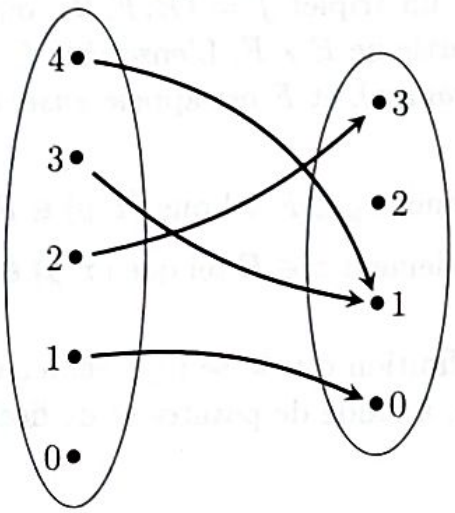
On peut maintenant définir la notion d'application.

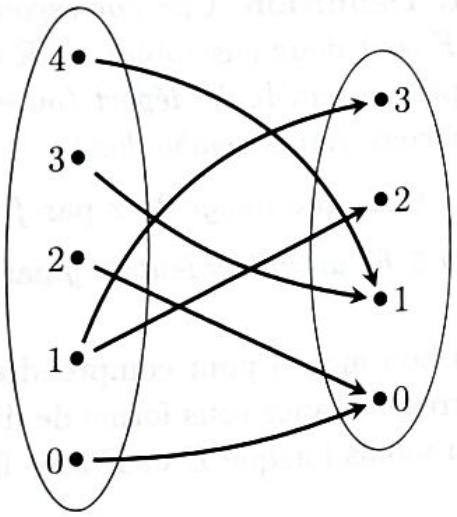
1.3. Définition. Une *application* est une correspondance $f = (E, F, R)$ telle que tout élément de E possède une unique image dans F. Autrement dit, une application est un triplet $f = (E, F, R)$, où E et F sont deux ensembles, et R est une partie de $E \times F$, vérifiant la propriété suivante : pour tout $x \in E$, il existe un unique $y \in F$ tel que $(x, y) \in R$.

Cet élément $y \in F$ est alors noté $f(x)$. C'est l'unique image de x par f. On utilise alors la notation $f : E \longrightarrow F$ ou encore

$$\begin{aligned} f\colon E &\longrightarrow F \\ x &\longmapsto f(x). \end{aligned}$$

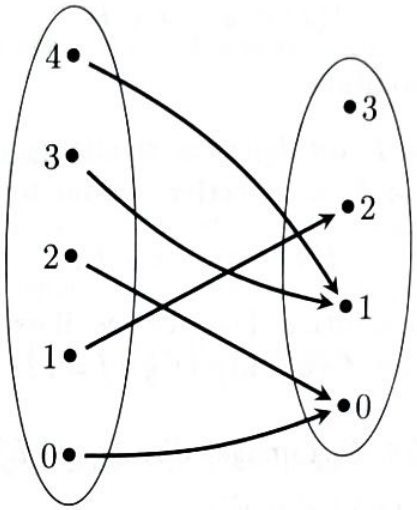
1.4. Exemple. La correspondance de gauche n'est pas une application, car $0 \in E$ n'a pas d'image dans F.





De même, la correspondance de droite n'est pas non plus une application, car $1 \in E$ admet deux images.

En revanche, la correspondance



est une application. On a

$$f(0) = 0, \quad f(1) = 2, \quad f(2) = 0, \quad f(3) = 1 \quad \text{et} \quad f(4) = 1.$$

1.5. Remarques

(1) Si $f : E \longrightarrow F$ est une application, par définition, **tout** élément de E doit avoir une image. Ainsi, nous excluons la notion de domaine de définition. Par exemple, la correspondance notée abusivement

$$\begin{aligned} \mathbb{R} &\longrightarrow \mathbb{R} \\ x &\longmapsto \frac{1}{x} \end{aligned}$$

n'est pas une application, car 0 n'a pas d'image.

(2) Si $f : E \longrightarrow F$ est une application, tout élément $x \in E$ a une unique image, mais un élément $y \in F$ peut avoir plusieurs antécédents, ou n'en avoir aucun.

1.6. Exemple. Considérons l'application

$$\begin{aligned} f : [-1, 1] &\longrightarrow \mathbb{R} \\ x &\longmapsto x^2. \end{aligned}$$

Alors, $y = -1 \in \mathbb{R}$ n'a pas d'antécédents par f, alors que $y = 1$ a deux antécédents, qui sont 1 et -1.

1.7. Définition. On dit que $f : E \longrightarrow F$ est *injective* si pour tous $x, x' \in E$, on a

$$f(x) = f(x') \quad \Longrightarrow \quad x = x'.$$

Autrement dit, $f : E \longrightarrow F$ est injective si pour tout $y \in F$, l'équation

$$f(x) = y, \quad x \in E,$$

possède *au plus* une solution.

On dit que $f : E \longrightarrow F$ est *surjective* si pour tout $y \in F$, l'équation

$$f(x) = y, \quad x \in E,$$

possède *au moins* une solution.

On dit que $f : E \longrightarrow F$ est *bijective* si elle est injective et surjective. Autrement dit, $f : E \longrightarrow F$ est bijective si pour tout $y \in F$, l'équation

$$f(x) = y, \quad x \in E,$$

possède *exactement une* solution. Dans ce cas, il existe une unique application [1] $g : F \longrightarrow E$ vérifiant $f \circ g = \mathrm{Id}_F$ et $g \circ f = \mathrm{Id}_E$. On la note f^{-1}.

1.8. Définition. Si $E' \subset E$, *l'image (directe)* de E' par f est l'ensemble

$$\begin{aligned} f(E') &= \{f(x') \mid x' \in E'\} \\ &= \{y \in F \mid \text{ il existe } x' \in E' \text{ tel que } y = f(x')\}. \end{aligned}$$

L'image de f est l'ensemble $f(E)$. On le note aussi $\mathrm{Im}(f)$.

Si $F' \subset F$, *l'image réciproque* de F' par f est l'ensemble

$$f^{-1}(F') = \{x \in E \mid f(x) \in F'\}.$$

Attention ! La notation est trompeuse, car f n'est pas nécessairement bijective. On trouve parfois la notation $f^{\langle -1 \rangle}(F')$ pour éviter la confusion.

1.9. Remarques

(1) On a $\mathrm{Im}(f) = F$ si, et seulement si, f est surjective.

(2) Si $f : E \longrightarrow F$ est une application, l'application

$$\begin{aligned} E &\longrightarrow \mathrm{Im}(f) \\ x &\longmapsto f(x) \end{aligned}$$

est surjective. En particulier, si une application $f : E \longrightarrow F$ est injective, elle induit une bijection de E sur son image.

2. Relations d'équivalence

La notion de relation d'équivalence sur un ensemble permet de mettre en relation des éléments ayant une propriété commune.

Cela permet alors regrouper ces éléments par « paquets », définissant ainsi la notion de classe d'équivalence, puis de construire un nouvel ensemble en « identifiant » les éléments d'un même paquet à un seul et même élément. On aboutit alors à la notion d'ensemble quotient.

1. Pour tout $y \in F$, $g(y)$ est simplement l'unique solution $x \in E$ de l'équation $f(x) = y$.

2.1. Définition. Soit E un ensemble. Une *relation* sur E est une partie R de $E \times E$. On note $x \sim_R y$ si $(x, y) \in R$.

Une *relation d'équivalence sur* E est une relation R satisfaisant les propriétés suivantes :

(1) réflexivité : pour tout $x \in E$, $x \sim_R x$;

(2) symétrie : pour tous x, $y \in E$, si $x \sim_R y$, alors $y \sim_R x$;

(3) transitivité : pour tous x, y, $z \in E$, si $x \sim_R y$ et $y \sim_R z$, alors $x \sim_R z$.

On confondra souvent R et $\sim_R$, et l'on parlera plutôt de la relation d'équivalence $\sim_R$ ou $\sim$.

Si $x \in E$, *la classe d'équivalence de* x est l'ensemble

$$\overline{x} = \{y \in E \mid y \sim_R x\}.$$

C'est le sous-ensemble des éléments en relation avec x. En particulier, $x \in \overline{x}$ pour tout $x \in E$.

2.2. Remarque. Si E est vide, alors R est vide. Il n'y a alors qu'une classe d'équivalence, qui est la classe vide. Si E est non vide, alors R est non vide (puisque R contient (x, x) pour tout $x \in E$). Dans ce cas, chaque classe d'équivalence est non vide.

2.3. Exemples

(1) La relation d'égalité est une relation d'équivalence sur tout ensemble E, associée à la partie $R = \{(x, x) \mid x \in E\}$. De plus, pour tout $x \in E$, on a

$$\overline{x} = \{x\}.$$

(2) La relation « avoir la même parité » est une relation d'équivalence sur l'ensemble $\mathbb{Z}$. Il y a deux classes d'équivalence : l'ensemble P des entiers pairs, et l'ensemble I des entiers impairs.

2.4. Définition. L'ensemble $E/_\sim$ dont les éléments sont les diverses classes d'équivalence de E pour la relation $\sim$ est appelé *ensemble quotient* de E par la relation d'équivalence $\sim$.

Attention ! Les éléments de $E/_\sim$ sont des sous-ensembles de E.

2.5. Exemple. Si $\mathbb{Z}$ est muni de la relation d'équivalence « avoir même parité », alors l'ensemble quotient a deux éléments, qui sont P et I.

2.6. Proposition. *Soit E un ensemble non vide, que l'on suppose muni d'une relation d'équivalence $\sim$.*
Alors :

(1) *pour tous $x, y \in E$, on a $x \sim y \iff \overline{x} = \overline{y}$;*

(2) *deux classes d'équivalence sont soit disjointes, soit égales ;*

(3) *les classes d'équivalence forment une partition de E. Autrement dit, les classes d'équivalence sont non vides, deux à deux disjointes et de réunion E.*

Démonstration. Si $\overline{x} = \overline{y}$, alors $x \in \overline{x} = \overline{y}$, et donc $x \sim y$. Inversement, supposons que $x \sim y$. Soit $z \in \overline{x}$. Alors $z \sim x$, et puisque $x \sim y$, on a $z \sim y$. Ainsi, $z \in \overline{y}$, et donc $\overline{x} \subset \overline{y}$. Mais puisque $x \sim y$, on a $y \sim x$ et le même raisonnement montre que $\overline{y} \subset \overline{x}$, et donc $\overline{x} = \overline{y}$.

Montrons maintenant le deuxième point. Soient $\overline{x}$ et $\overline{y}$ deux classes d'équivalence et supposons qu'elles ne soient pas disjointes. Autrement dit, il existe $z \in E$ tel que $z \in \overline{x}$ et $z \in \overline{y}$. Par définition, on a $x \sim z$ et $y \sim z$. On a alors $x \sim z$ et $z \sim y$, et par suite $x \sim y$. Le premier point implique alors que $\overline{x} = \overline{y}$.

Enfin, pour montrer le dernier point, il reste à montrer que les classes d'équivalence sont non vides, de réunion E, le point (2) montrant que ces classes sont deux à deux disjointes. Mais c'est clair puisque l'on a $x \in \overline{x}$ pour tout $x \in E$. □

2.7. Corollaire. *Soit $\omega \in E/\sim$ une classe d'équivalence pour la relation d'équivalence $\sim$. Alors, pour tout $x \in \omega$, on a $\omega = \overline{x}$.*

Démonstration. Par définition d'une classe d'équivalence, on a $\omega = \overline{y}$ pour un certain $y \in E$. Si $x \in \omega = \overline{y}$, alors $x \sim y$, et donc $\overline{x} = \overline{y} = \omega$ par la proposition 2.6 (1). □

2.8. Définition. Soit $\omega \in E/\sim$ une classe d'équivalence pour la relation d'équivalence $\sim$. Un élément $x \in E$ tel que $\omega = \overline{x}$ est appelé un *représentant* de ω. Par le corollaire précédent, $x \in E$ est un représentant de ω si, et seulement si, $x \in \omega$.

2.9. Exemple. Si $E = \mathbb{Z}$ est muni de la relation d'équivalence « avoir même parité », 0 et 2 sont des représentants de l'ensemble P des entiers pairs.

2.10. Définition. Soit E un ensemble non vide muni d'une relation d'équivalence $\sim$. L'application surjective

$$\begin{aligned} \pi : E &\longrightarrow E/_\sim \\ x &\longmapsto \overline{x} \end{aligned}$$

est appelée la *projection canonique.*

On veut maintenant caractériser l'ensemble des applications

$$\varphi : E/_\sim \longrightarrow F,$$

où F est un ensemble.

Remarquons tout d'abord que si

$$\varphi : E/_\sim \longrightarrow F$$

est une application, alors l'application $\varphi \circ \pi : E \longrightarrow F$ est constante sur les classes d'équivalence.

En effet, si $x \sim y$, alors $\overline{x} = \overline{y}$, et donc

$$(\varphi \circ \pi)(x) = \varphi\big(\pi(x)\big) = \varphi(\overline{x}) = \varphi(\overline{y}) = \varphi\big(\pi(y)\big) = (\varphi \circ \pi)(y).$$

Réciproquement, on a le résultat suivant.

2.11. Proposition. *Soit E un ensemble non vide, muni d'une relation d'équivalence $\sim$, soit F un ensemble non vide et soit $f : E \longrightarrow F$ une application. On suppose que f est constante sur les classes d'équivalence, c'est-à-dire*

$$\text{pour tous } x, y \in E, \quad x \sim y \implies f(x) = f(y).$$

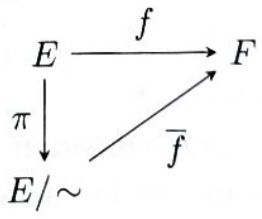
Alors, il existe une unique application

$$\overline{f} : E/\sim \longrightarrow F$$

telle que $f = \overline{f} \circ \pi$. Elle est définie par

$$\overline{f}(\omega) = f(x), \qquad \text{pour tout } \omega \in E/_\sim,$$

où $x \in E$ est un représentant arbitraire de ω. En particulier, on a



$$\overline{f}(\overline{x}) = f(x), \qquad \text{pour tout } x \in E.$$

En outre, pour tout ensemble non vide F, il existe une correspondance bijective entre l'ensemble des applications $\varphi : E/_\sim \longrightarrow F$ et l'ensemble des applications $f : E \longrightarrow F$ qui sont constantes sur les classes d'équivalence. Cette bijection est donnée par

$$\begin{aligned} \varphi &\longmapsto \varphi \circ \pi \\ \overline{f} &\longmapsfrom f. \end{aligned}$$

Démonstration. Gardons les notations de l'énoncé. Supposons qu'une telle application $\overline{f}$ existe et soit $\omega \in E/\sim$. Si $x \in E$ est un représentant de ω, on a $\omega = \overline{x} = \pi(x)$, et ainsi

$$\overline{f}(\omega) = \overline{f}\big(\pi(x)\big) = (\overline{f} \circ \pi)(x) = f(x).$$

Cela montre que si $\overline{f}$ existe, elle est définie par

$$\overline{f}(\omega) = f(x), \quad \text{pour tout } \omega \in E/\sim,$$

où $x \in E$ est un représentant arbitraire de ω. En particulier, une telle application est alors unique.

Le problème est de vérifier que réciproquement cette formule définit bien une application.

Autrement dit, il faut vérifier que $\overline{f}(\omega)$ ne dépend pas du représentant de la classe d'équivalence choisi. Soient $x, y \in E$ deux représentants de ω. On a donc $x, y \in \omega$. On doit montrer alors que $f(x) = f(y)$. Mais c'est exactement l'hypothèse faite sur f.

Cela démontre l'existence de l'application $\overline{f}$. De plus, pour tout $x \in E$, x est un représentant de $\overline{x}$, et l'on a donc

$$\overline{f}(\overline{x}) = f(x), \quad \text{pour tout } x \in E.$$

Pour achever la démonstration du théorème, il suffit de vérifier que les deux flèches

$$\begin{array}{ccc} \varphi & \longmapsto & \varphi \circ \pi \\ \overline{f} & \longleftarrow\!\shortmid & f \end{array}$$

sont inverses l'une de l'autre.

Soit $\varphi : E/\sim \longrightarrow F$ une application. Alors, l'application $\overline{\varphi \circ \pi}$ est l'unique application $\psi : E/\sim \longrightarrow F$ telle que $\psi \circ \pi = \varphi \circ \pi$. Or, φ est une telle application. On en déduit alors que $\overline{\varphi \circ \pi} = \varphi$. Inversement, si $f : E \longrightarrow F$ est constante sur les classes d'équivalence, on a $\overline{f} \circ \pi = f$ par définition. Cela achève la démonstration. □

2.12. Remarque. L'hypothèse sur f est indispensable !

Reprenons l'exemple de la relation d'équivalence « avoir la même parité » sur $\mathbb{Z}$ et soit $f = \mathrm{Id}_{\mathbb{Z}}$. Alors, f n'est pas constante sur les classes d'équivalence, et l'application $\overline{f}$ précédente n'existe pas. En effet, dans le cas contraire, on aurait $\overline{f}(P) = f(0) = 0$, mais aussi $\overline{f}(P) = f(2) = 2$, car les entiers 0 et 2 sont deux représentants de la classe d'équivalence formée par les entiers pairs. On obtiendrait alors l'égalité $0 = 2 \in \mathbb{Z}$.

3. Lois internes, structures algébriques

Les structures sont les armes des mathématiciens[2].

Dans tout ce qui suit, E est un ensemble non vide.

3.1. Définition. Une *loi interne* $*$ sur E est une application

$$E \times E \longrightarrow E$$
$$(x, y) \longmapsto x * y.$$

On dit que la loi $*$:

(1) est *associative* si pour tous x, y, $z \in E$, on a $(x * y) * z = x * (y * z)$;

(2) est *commutative* si pour tous x, $y \in E$, on a $x * y = y * x$.

Si $\perp$ est une autre loi interne sur E, on dit que la loi $*$ est *distributive* par rapport à $\perp$ si, pour tous x, y, $z \in E$, on a

$$(x \perp y) * z = (x * z) \perp (y * z) \quad \text{et} \quad x * (y \perp z) = (x * y) \perp (x * z).$$

3.2. Exemple. L'addition et le produit sont des lois associatives sur $\mathbb{N}$, ainsi que sur $\mathbb{Z}$. De plus, le produit est distributif par rapport à l'addition.

3.3. Définition. Un *élément neutre* pour la loi $*$ est un élément $e \in E$ tel que, pour tout $x \in E$, on a $x * e = e * x = x$.

3.4. Exemple. L'addition sur $\mathbb{N}$ possède un élément neutre, à savoir 0, et le produit sur $\mathbb{N}$ possède un élément neutre, à savoir 1.

3.5. Lemme. *Soit E un ensemble non vide, muni d'une loi interne $*$. Si la loi $*$ possède un élément neutre, celui-ci est unique.*

Démonstration. Supposons que $*$ possède deux éléments neutres e_1, $e_2 \in E$. Puisque e_1 est un élément neutre, on a $e_1 * e_2 = e_2$. Mais, e_2 étant aussi un élément neutre, on a également $e_1 * e_2 = e_1$. On en déduit que $e_1 = e_2$, d'où le résultat. □

3.6. Définition. Soit E un ensemble non vide, muni d'une loi $*$, et possédant un élément neutre e (nécessairement unique d'après le lemme précédent). Soit $x \in E$. Un *symétrique* de x est un élément $x' \in E$ tel que

$$x * x' = x' * x = e.$$

2. Nicolas Bourbaki.

3.7. Exemple. Tout $m \in \mathbb{Z}$ possède un symétrique dans $\mathbb{Z}$ pour la loi d'addition, à savoir $-m$. En revanche, le seul entier naturel possédant un symétrique pour l'addition dans $\mathbb{N}$ est 0.

3.8. Lemme. *Soit E un ensemble non vide, muni d'une loi interne $*$. On suppose que $*$ est associative et possède un élément neutre e. Si $x \in E$ possède un symétrique, celui-ci est unique.*

Démonstration. Supposons que $x \in E$ possède deux symétriques $x'_1, x'_2 \in E$. Alors, on a

$$x'_1 = x'_1 * e = x'_1 * (x * x'_2) = (x'_1 * x) * x_2 = e * x'_2 = x'_2,$$

d'où le résultat. □

On peut maintenant définir la notion de groupe.

3.9. Définition. Un *groupe* est un couple $(G, *)$, où G est un ensemble non vide, et $* : G \times G \longrightarrow G$ est une loi interne associative, possédant un élément neutre $1_G \in G$, et telle que tout élément $x \in G$ possède un symétrique.

D'après les résultats du paragraphe précédent, le neutre est unique, et un symétrique de $x \in G$ est unique. On le note x^{-1}, et on l'appelle *l'inverse* de l'élément x.

Un groupe $(G, *)$ est dit *abélien* (ou commutatif) si la loi $*$ est commutative.

Les groupes seront étudiés de manière approfondie dans la partie suivante. Aussi, nous nous bornerons pour l'instant à donner deux exemples.

3.10. Exemples

(1) L'ensemble $\mathbb{Z}$ des entiers relatifs, muni de son addition, est un groupe abélien, de neutre 0.

(2) L'ensemble $\mathbb{R}^\times = \mathbb{R}\backslash\{0\}$, muni du produit usuel, est un groupe abélien, de neutre 1.

3.11. Remarques

(1) Dans les énoncés, on omettra de préciser la loi $*$, et l'on parlera du groupe G, plutôt que du groupe $(G, *)$.
D'ailleurs, si $x, y \in G$, on écrira xy plutôt que $x * y$.

(2) Lorsque la loi de G est commutative, on la note souvent (mais pas toujours) « + ». Dans ce cas, le neutre est noté 0_G, et l'inverse de x est plutôt noté $-x$, et appelé « opposé de x ».

Nous allons maintenant définir la notion de sous-groupe.

3.12. Définition. Soit G un groupe, de neutre 1_G. Un *sous-groupe* de G est un sous-ensemble H de G vérifiant les propriétés suivantes :

(1) on a $1_G \in H$;

(2) pour tous $x, y \in H$, on a $xy \in H$;

(3) pour tout $x \in H$, on a $x^{-1} \in H$.

3.13. Exemples

(1) L'ensemble $\mathbb{Z}$ est un sous-groupe de $(\mathbb{C}, +)$.

(2) L'ensemble $\{\pm 1\}$ est un sous-groupe de $(\mathbb{R}^\times, \cdot)$.

L'intérêt de cette notion est donné par le lemme suivant.

3.14. Lemme. *Soit G un groupe, et soit H un sous-groupe de G. Alors, la loi interne de G se restreint en une loi interne sur H, qui confère à H une structure de groupe.*

Démonstration. Le fait que la loi de G se restreigne en une loi interne sur l'ensemble H provient de l'axiome (2) de la définition de sous-groupe. Puisque G est un groupe, sa loi est associative. Sa restriction à H l'est donc aussi. De plus, 1_G étant un élément neutre pour la loi de G, c'est aussi un élément neutre pour sa restriction à H, car $1_G \in H$ par hypothèse sur H. Enfin, si $x \in H$, son symétrique x^{-1} est dans H. Mais alors, x^{-1} est un inverse de x pour la loi de H, puisque $xx^{-1} = x^{-1}x = 1_G = 1_H$. □

Ce résultat montre que, pour vérifier qu'un ensemble muni d'une loi interne est un groupe, il vaut mieux l'identifier à un sous-groupe d'un groupe connu, puisque cela permet d'éviter de démontrer l'associativité (chose qui est somme toute assez pénible à faire, en général).

Maintenant que l'on a défini les groupes, on va ajouter une loi supplémentaire, et définir la notion d'anneau. Encore une fois, nous nous limiterons pour l'instant au strict minimum.

3.15. Définition. Un *anneau* est un triplet $(A, +, \cdot)$, où A est un ensemble non vide, $+ : A \times A \longrightarrow A$ et $\cdot : A \times A \longrightarrow A$ sont deux lois internes, vérifiant les propriétés suivantes :

(1) le couple $(A, +)$ est un groupe abélien ;

(2) la loi « $\cdot$ » est associative, et possède un élément neutre[3] ;

(3) la loi « $\cdot$ » est distributive par rapport à « $+$ ».

3. Selon certains auteurs, un anneau ne possède pas nécessairement un tel élément neutre, et appellent « anneau unitaire » ce que nous appelons un anneau. Néanmoins, notre convention est la plus répandue actuellement.

Les éléments neutres pour les lois « + » et « · » seront notés 0_A et 1_A respectivement.

On dit que l'anneau $(A, +, \cdot)$ est *commutatif* si « · » est commutative.

3.16. Exemples

(1) L'ensemble $\mathbb{Z}$ des entiers relatifs, muni de son addition et de son produit, est un anneau commutatif.

(2) L'ensemble $\mathrm{M}_n(\mathbb{R})$ des matrices carrées de taille n à coefficients réels, muni de l'addition et du produit de matrices, est un anneau, non commutatif si $n \geqslant 2$.

Nous aurons aussi besoin rapidement de la notion de corps.

3.17. Définition. Un *corps* est un ensemble K, muni de deux lois internes

$$\begin{array}{ccc} K \times K \longrightarrow K \\ (x, y) \longmapsto x + y \end{array} \quad \text{et} \quad \begin{array}{ccc} K \times K \longrightarrow K \\ (x, y) \longmapsto xy, \end{array}$$

appelées respectivement addition et multiplication, vérifiant les propriétés suivantes :

(1) le triplet $(K, +, \cdot)$ est un anneau commutatif ;

(2) tout élément *non nul* de K possède un symétrique par rapport à la loi « · » ;

(3) on a $0_K \neq 1_K$.

Le symétrique de $x \in K$ par rapport à l'addition est noté $-x$, et appelé *l'opposé* de x. Le symétrique de $x \in K \setminus \{0_K\}$ par rapport à la multiplication est noté x^{-1}, et appelé *l'inverse* de x.

On notera $K^\times = K \setminus \{0_K\}$.

3.18. Remarque. Avec cette définition, un corps est en particulier un anneau commutatif. Ce n'est pas toujours le cas dans les ouvrages. Nous préférons néanmoins réserver le terme « corps » pour désigner ce que d'autres auteurs appelleraient « corps commutatif », et plutôt parler d'anneau à division, ou de corps gauche, lorsque le groupe multiplicatif $(K \setminus \{0_K\}, \cdot)$ n'est pas abélien [4].

3.19. Exemple. Les ensembles $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$, munis de leurs lois usuelles, sont des corps. Des exemples de corps un peu plus exotiques seront introduits dans des chapitres ultérieurs.

4. La convention utilisée ici l'est d'ailleurs aussi chez nos amis anglo-saxons.

On finit ce chapitre par des considérations sur les relations d'ordre. Les résultats de ce paragraphe ne seront utilisés qu'à une ou deux reprises dans cet ouvrage, et peuvent être passés en première lecture.

4. Ensembles ordonnés

4.1. Définition. Soit E un ensemble, et soit « $\leqslant$ » une relation sur E. On dit alors que « $\leqslant$ » est une *relation d'ordre* si l'on a les propriétés suivantes :

(1) réflexivité : pour tout $x \in E$, on a $x \leqslant x$;

(2) antisymétrie : pour tous $x, y \in E$, $(x \leqslant y \text{ et } y \leqslant x) \Longrightarrow x = y$;

(3) transitivité : pour tous $x, y, z \in E$, $(x \leqslant y \text{ et } y \leqslant z) \Longrightarrow x \leqslant z$.

On note alors $x < y$ si $x \leqslant y$ et $x \neq y$.

On dit que la relation d'ordre « $\leqslant$ » est *totale*, ou que $(E, \leqslant)$ est totalement ordonné, si pour tous $x, y \in E$, on a $x \leqslant y$ ou $y \leqslant x$.

4.2. Exemples

(1) Soit E un ensemble. Alors, la relation d'inclusion est une relation d'ordre sur l'ensemble $\mathcal{P}(E)$ des parties de E. Cette relation d'ordre n'est pas totale dès que E possède au moins deux éléments.

(2) La relation d'ordre usuelle sur l'ensemble $\mathbb{N}$ est une relation d'ordre total.

4.3. Définition. Soit $(E, \leqslant)$ un ensemble ordonné.

On dit que $z \in E$ est *maximal* si pour tout $x \in E$, on a $z \leqslant x \implies x = z$. Autrement dit, $z \in E$ est maximal s'il n'existe aucun élément $x \in E$ tel que $z < x$.

On dit que $z \in E$ est un *maximum* (ou un *plus grand élément*) si $x \leqslant z$ pour tout $x \in E$.

4.4. Remarques

(1) Un maximum, s'il existe, est unique, et c'est aussi le seul élément maximal.

En effet, si $z, z' \in E$ sont deux maximums de E, on a $z' \leqslant z$ et $z \leqslant z'$ par définition d'un maximum, et donc $z' = z$, par définition d'une relation d'ordre.

De plus, si $z \in E$ est un maximum, et si $x \in E$ vérifie $z \leqslant x$, alors puisque $x \leqslant z$ par définition d'un maximum, on a $x = z$. Par ailleurs, si $z' \in E$ est maximal, et puisque $z' \leqslant z$ par définition d'un maximum, on a $z = z'$.

(2) En revanche, un élément maximal n'est pas nécessairement un maximum.
Par exemple, soit E l'ensemble des parties strictes de $X = \{0, 1, 2\}$, ordonné par l'inclusion. Alors, toute partie de X à deux éléments est un élément maximal, mais E ne possède pas d'élément maximum, car il n'existe pas de partie stricte de X qui contienne toutes les autres.

(3) Tout ensemble fini totalement ordonné possède un maximum.

4.5. Définition. Soit $(E, \leqslant)$ un ensemble ordonné, et soit F une partie non vide de E. On dit que $x \in E$ est un *majorant* de F si l'on a

$$y \leqslant x, \quad \text{pour tout } y \in F.$$

On dit que E est *inductif* si tout sous-ensemble de E totalement ordonné admet un majorant.

On a alors le résultat suivant.

4.6. Théorème. [Lemme de Zorn] *Tout ensemble inductif non vide possède un élément maximal.*

On se reportera à [42], théorème 1.6.20 pour une démonstration.

5. Exercices

E1. Soient E, F deux ensembles, soit $f : E \longrightarrow F$ une application, et soient enfin $(A_i)_{i\in I}$ une famille de parties de E et $(B_i)_{i\in I}$ une famille de parties de F.

a. Quelles sont les inclusions entre :

(i) $f\Big(\bigcup_{i\in I} A_i\Big)$ et $\bigcup_{i\in I} f(A_i)$;

(ii) $f\Big(\bigcap_{i\in I} A_i\Big)$ et $\bigcap_{i\in I} f(A_i)$;

(iii) $f^{-1}\Big(\bigcup_{i\in I} B_i\Big)$ et $\bigcup_{i\in I} f^{-1}(B_i)$;

(iv) $f^{-1}\Big(\bigcap_{i\in I} B_i\Big)$ et $\bigcap_{i\in I} f^{-1}(B_i)$?

b. Que peut-on dire si l'on suppose f injective ou surjective ?

E2. Dans les cas suivants, peut-on définir une application $f : \mathbb{R} \longrightarrow \mathbb{R}$ par

(i) $f(t^3) = t$, pour tout $t \in \mathbb{R}$;

(ii) $f(t^2) = t$, pour tout $t \in \mathbb{R}$;

(iii) $f(a+b) = a - b$, pour tous $a, b \in \mathbb{R}$;

(iv) $f\left(\frac{a}{b}\right) = a - b$, pour tous $a, b \in \mathbb{R}$, $b \neq 0$?

E3.

a. Considérons l'application

$$\begin{array}{rccc} f : & \mathbb{R} & \longrightarrow & \mathbb{U} = \{z \in \mathbb{C} \mid |z| = 1\} \\ & t & \longmapsto & e^{2i\pi t}. \end{array}$$

Existe-t-il une application $g : \mathbb{U} \longrightarrow \mathbb{R}$ telle que $f \circ g = \mathrm{Id}_{\mathbb{U}}$?

b. Soit $f : E \longrightarrow F$ une application entre deux ensembles E et F.

(i) Donner une condition nécessaire et suffisante pour qu'il existe une application $g : F \longrightarrow E$ telle que $f \circ g = \mathrm{Id}_F$.

(ii) Donner une condition nécessaire et suffisante pour qu'il existe une application $h : E \longrightarrow F$ telle que $h \circ f = \mathrm{Id}_E$.

E4. La relation binaire $\sim$ sur l'ensemble E est-elle une relation d'équivalence dans les cas suivants ?

a. $E = \mathbb{R}$ et $x \sim y$ si

(i) $x - y \in 3\mathbb{Z}$;

(ii) $x + y \in \mathbb{Z}$;

(iii) $|x - y| \leqslant 2$;

(iv) $x \leqslant y$.

b. $E = \mathbb{R}^\times = \mathbb{R} \setminus \{0\}$ et $x \sim y$ si $\frac{x}{y} \in \mathbb{Q}$.

c. E est l'ensemble des droites du plan, et $d_1 \sim d_2$ si d_1 est parallèle (ou confondue) à d_2.

d. E est l'ensemble des droites du plan, et $d_1 \sim d_2$ si d_1 est perpendiculaire à d_2.

E5. Que pensez-vous du raisonnement suivant : « Soit $\sim$ une relation binaire symétrique et transitive sur un ensemble E. Par symétrie de $\sim$, on a, pour tout $(x, y) \in E \times E$, $x \sim y \Longrightarrow y \sim x$. Par transitivité de $\sim$, on a $(x \sim y$ et $y \sim x) \Longrightarrow x \sim x$. Ainsi, $\sim$ est réflexive » ?

E6. Une relation d'équivalence $\sim$ sur un ensemble E est par définition une relation réflexive, symétrique et transitive. Donner des exemples de relations binaires $\sim$ sur un ensemble E qui vérifient deux de ces propriétés mais pas la troisième.

E7. Montrer que la relation binaire $\sim$ sur l'ensemble E est une relation d'équivalence dans les cas suivants :

a. $E = \mathbb{R}^n$ et, pour tous $x, y \in E$, $x \sim y$ si $y - x \in \mathbb{Z}^n$.

b. $E = \{x \in \mathbb{R}^n \mid ||x||_2 = 1\}$ et, pour tous x, $y \in E$, $x \sim y$ si $y = \pm x$.

c. $E = \mathbb{R}^n \setminus \{(0, \ldots, 0)\}$ (où $n \geqslant 1$) et, pour tous x, $y \in E$, $x \sim y$ s'il existe $\lambda \in \mathbb{R}^\times$, $y = \lambda x$.

L'ensemble quotient $E/\sim$ correspondant est appelé *espace projectif réel de dimension* $n-1$ et noté $\mathbb{P}^{n-1}_{\mathbb{R}}$.

d. $E = \mathcal{F}(\mathbb{R}, \mathbb{R})$ et, pour tous f, $g \in E$, $f \sim g$ si $f(0) = g(0)$.

e. $E = \mathcal{F}(\mathbb{R}^\times, \mathbb{R})$ et, pour tous f, $g \in E$, $f \sim g$ si $f - g = o(x)$.

f. $E = \mathbb{Z} \times (\mathbb{Z} \setminus \{0\})$ et, pour tous (a, b), $(c, d) \in E$, $(a, b) \sim (c, d)$ si l'on a l'égalité $ad - bc = 0$. On n'utilisera que des calculs dans $\mathbb{Z}$.

Quel est l'ensemble quotient dans ce cas ?

g. Soit E l'ensemble des couples de points du plan. Pour tous couples de points (A, B), $(C, D) \in E$, on dit que $(A, B) \sim (C, D)$ si les segments $[AD]$ et $[BC]$ ont même milieu. On admet que c'est une relation d'équivalence.

Que représente la classe d'équivalence de (A, B) ?

E8. Soit $E = \mathbb{N} \times \mathbb{N}$ et soit $\sim$ la relation binaire sur E définie par

$$(a, b) \sim (c, d) \iff a + d = b + c.$$

Dans cet exercice, on se limitera à des calculs dans $\mathbb{N}$.

a. Montrer que $\sim$ est une relation d'équivalence.

b. Montrer que toute classe d'équivalence est représentée de manière unique par un élément de la forme $(n, 0)$, $n \geqslant 0$ ou $(0, m)$, $m \geqslant 1$.

c. On note $[(a, b)]$ la classe d'équivalence de (a, b). Montrer que l'application

$$\begin{aligned} \mathbb{N} &\longrightarrow E/\sim \\ n &\longmapsto [(n, 0)] \end{aligned}$$

est injective.

d. Montrer que si $(a, b) \sim (a', b')$ et $(c, d) \sim (c', d')$, alors

$$(a + c,\, b + d) \sim (a' + c',\, b' + d').$$

e. Montrer que la loi interne

$$\begin{aligned} E/\sim \times E/\sim &\longrightarrow E/\sim \\ \big([(a, b)],\, [(c, d)]\big) &\longmapsto [(a, b)] + [(c, d)] = [(a + c,\, b + d)] \end{aligned}$$

est bien définie, et que $(E/\sim, +)$ est un groupe abélien.

f. À votre avis, quel est ce groupe ?

E9. Si E est un ensemble non vide E, on rappelle qu'une *partition* de E est une famille $(E_i)_{i\in I}$ de sous-ensembles de E, non vides, deux à deux disjoints, dont la réunion est E.

Étant donné une partition $(E_i)_{i\in I}$ d'un ensemble non vide E, montrer qu'il existe une unique relation d'équivalence sur E dont les classes d'équivalence sont les E_i, $i \in I$.

E10. Soit $f : E \longrightarrow E'$ une application. Soit $\sim$ une relation d'équivalence sur E et soit $\sim'$ une relation d'équivalence sur E'. On suppose que

$$\text{pour tous } x, y \in E, \quad x \sim y \implies f(x) \sim' f(y).$$

Montrer qu'il existe une unique application $\varphi : E/_\sim \longrightarrow E'/_{\sim'}$ qui envoie la classe d'équivalence de x sur la classe d'équivalence de $f(x)$.

E11. On définit la relation $\sim$ sur $\mathbb{R}$ par

$$t \sim t' \iff t' - t \in \mathbb{Z}.$$

a. Vérifier que c'est une relation d'équivalence.

L'ensemble quotient est noté $\mathbb{R}/\mathbb{Z}$.

b. Montrer que toute classe d'équivalence est représentée par un unique élément de $[0, 1[$.

c. Montrer que $\mathbb{R}/\mathbb{Z}$ est en bijection avec $\mathbb{U}$, en utilisant l'application

$$\begin{aligned} \mathbb{R} &\longrightarrow \mathbb{U} \\ t &\longmapsto e^{2i\pi t}. \end{aligned}$$

E12. Soient $f : E \longrightarrow F$ une application et $\sim_F$ une relation d'équivalence sur F.

a. Montrer que la relation binaire $\sim_{E;f}$ sur E définie par

$$x \sim_{E;f} y \iff f(x) \sim_F f(y)$$

est une relation d'équivalence.

b. Décrire la classe d'équivalence de $x \in E$ pour la relation $\sim_{E;f}$ dans le cas où $\sim_F$ est l'égalité. On note alors simplement $\sim_f$ pour $\sim_{E;f}$.

c. Montrer que la relation binaire $\sim$ sur $\mathbb{R}^2$ définie par

$$(x, y) \sim (x', y') \quad \text{si} \quad x + y = x' + y'$$

est une relation d'équivalence, et décrire la classe de $(0, 1)$, ainsi que l'ensemble quotient.

E13.

a. Montrer que l'ensemble quotient du point **b** de l'exercice 7 est en bijection avec $\mathbb{P}^{n-1}_{\mathbb{R}}$ (défini au point **c** de ce même exercice).

b. Donner un ensemble « simple » en bijection avec l'ensemble quotient correspondant à l'exemple **d** de l'exercice 7.

Indication. Dans chaque cas, on introduira une fonction convenable, et l'on montrera qu'elle passe au quotient en une bijection.

Chapitre II

Arithmétique dans $\mathbb{Z}$

Dans ce chapitre, on rappelle les propriétés arithmétiques de $\mathbb{Z}$.

On suppose connus les ensembles $\mathbb{N}$ et $\mathbb{Z}$, ainsi que les propriétés de l'addition et de la multiplication, y compris par rapport à la relation d'ordre naturelle $\leqslant$.

Enfin, on suppose connu le fait suivant : toute partie non vide de $\mathbb{N}$ possède un plus petit élément.

Notation. Si a, $b \in \mathbb{N}$, $a \leqslant b$, on note $[\![a, b]\!]$ l'ensemble

$$[\![a, b]\!] = \{m \in \mathbb{N} \mid a \leqslant m \leqslant b\}.$$

1. Divisibilité dans $\mathbb{Z}$

On commence par introduire la notion de divisibilité.

1.1. Définition. Soient a, $b \in \mathbb{Z}$. On dit que a *divise* b s'il existe $c \in \mathbb{Z}$ tel que $b = ac$. Dans ce cas, on dit que a est un *diviseur* de b, ou que b est un *multiple* de a.

On note cela $a \mid b$.

1.2. Remarques

(1) Si $a \mid b$ et $b \mid a$, alors $b = \pm a$.
En effet, les hypothèses impliquent qu'il existe u, $v \in \mathbb{Z}$ tels que $b = ua$ et $a = vb$. Mais alors, $b = uvb$, soit

$$(1 - uv)b = 0.$$

Si $b = 0$, alors $a = vb = 0$. Si $b \neq 0$, on a $1 - uv = 0$, d'où $uv = 1$. On en déduit que u et v sont non nuls, de même signe.

Si u et v sont positifs, alors $u, v \geqslant 1$. On a alors $1 \leqslant u \leqslant uv = 1$, et donc $u = 1$, d'où $b = a$. Si u et v sont négatifs, $-u$ et $-v$ sont positifs, et le cas précédent montre que $u = -1$, soit $b = -a$.

(2) La relation de divisibilité est une relation d'ordre (partiel) sur $\mathbb{N}$.

En effet, la réflexivité et la transitivité découlent immédiatement des définitions, et l'antisymétrie découle du point (1).

(3) Si $a \mid b$, alors $a \mid ua + bv$, pour tous u, $v \in \mathbb{Z}$.

(4) Si $a \mid b$ et $b \neq 0$, on a $|a| \leqslant |b|$.

En effet, si $a \mid b$, il existe $c \in \mathbb{Z}$ tel que $b = ac$. Comme b est non nul, c est non nul. En particulier, $|c| \geqslant 1$, et l'on a

$$|a| \leqslant |a| \times |c|,$$

c'est-à-dire $|a| \leqslant |b|$.

Nous allons maintenant rappeler les notions de pgcd et de ppcm.

1.3. Définition. Soient a, $b \in \mathbb{Z}$. On dit que $d \in \mathbb{Z}$ est un *pgcd* de a et b si d vérifie les propriétés suivantes :

(1) on a $d \mid a$ et $d \mid b$;

(2) pour tout $c \in \mathbb{Z}$ tel que $c \mid a$ et $c \mid b$, on a $c \mid d$.

On dit que $m \in \mathbb{Z}$ est un *ppcm* de a et b si m vérifie les propriétés suivantes :

(1) on a $a \mid m$ et $b \mid m$;

(2) pour tout $c \in \mathbb{Z}$ tel que $a \mid c$ et $b \mid c$, on a $m \mid c$.

1.4. Remarques

(1) Si $a = 0$, un pgcd de a et b est b.

En effet, b divise bien à la fois 0 et b. Si maintenant $c \in \mathbb{Z}$ divise à la fois 0 et b, il divise b, d'où le résultat.

En particulier, si $a = b = 0$, un pgcd de a et b est 0.

(2) Si $(a, b) \neq (0, 0)$, un pgcd de a et b (lorsqu'il existe) est non nul. En effet, si par exemple $a \neq 0$, on a $a = da'$, $a' \in \mathbb{Z}$. Comme $a \neq 0$, on a nécessairement $d \neq 0$.

(3) Si $a = 0$ ou $b = 0$, un ppcm de a et b est 0. En effet, 0 est un multiple commun à a et b. Si maintenant, c est un multiple commun à a et b, alors $c = 0$. Mais alors, c est un multiple de 0.

(4) Si a et b sont non nuls, un ppcm m de a et b (s'il existe) est nécessairement non nul, puisqu'il divise ab, qui est non nul.

(5) Un pgcd et un ppcm de a et b (s'ils existent) sont uniques au signe près.

En effet, soient d et d' deux pgcd de a et b. Alors, $d' \mid a$ et $d' \mid b$, par définition de d', et donc $d' \mid d$ par définition de d. De même, $d \mid d'$, et par la remarque 1.2 (1), on obtient $d' = \pm d$.

Un raisonnement similaire montre qu'un ppcm est unique au signe près.

Nous allons maintenant voir qu'un pgcd et un ppcm existent toujours. Pour cela, nous aurons besoin de la division euclidienne.

2. Division euclidienne, algorithme d'Euclide

Avant d'énoncer et de démontrer le théorème de la division euclidienne, nous avons besoin d'un résultat intermédiaire.

2.1. Lemme. *Toute partie non vide majorée de $\mathbb{Z}$ a un plus grand élément.*

Démonstration. Soit E une partie non vide majorée de $\mathbb{Z}$. Soit $M \in \mathbb{Z}$ un majorant de E, et soit

$$F = \{M - x \mid x \in E\}.$$

Par définition de M, l'ensemble F est une partie de $\mathbb{N}$, non vide puisque E est non vide. Par conséquent, F possède un plus petit élément y_0.

Soit $x_0 \in E$ tel que $y_0 = M - x_0$. Montrons que x_0 est le plus grand élément de E. Soit $x \in E$. Alors, $M - x \in F$, et ainsi $M - x \geqslant y_0$ par définition de y_0. Puisque $y_0 = M - x_0$, on déduit immédiatement que $x \leqslant x_0$. Ainsi, x_0 est un majorant de E, et comme $x_0 \in E$, c'est aussi le plus grand élément de E. Cela achève la démonstration. □

Nous sommes maintenant prêts à démontrer le résultat principal de ce paragraphe.

2.2. Théorème (Division euclidienne). *Soient a, $b \in \mathbb{Z}$, avec b non nul. Alors, il existe q, $r \in \mathbb{Z}$ tels que*

(1) $a = bq + r$;

(2) $0 \leqslant r < |b|$.

De plus, les entiers q et r sont uniques.

Démonstration. Commençons par démontrer l'existence de q et r.

Supposons d'abord que $b > 0$. Posons $E = \{m \in \mathbb{Z} \mid mb \leqslant a\}$. Montrons que E est non vide et majorée.

Si $a \geqslant 0$, on a $0 \in E$. De plus, a est un majorant de E. En effet, supposons qu'il existe $m \in E$ tel que $m > a$. Alors, $m > 0$, et l'on aurait

$$a < m \leqslant mb \leqslant a,$$

d'où une contradiction. On a donc $m \leqslant a$ pour tout $m \in E$.

Si $a < 0$, on a $a \in E$, puisque $b \geqslant 1$ implique alors que $ab \leqslant a$. De plus, pour tout $m \in E$, 0 est un majorant de E. En effet, si E contenait un élément $m > 0$, on obtiendrait la contradiction

$$0 < mb \leqslant a < 0.$$

Par le lemme 2.1, la sous-ensemble E possède un plus grand élément q. On a donc $q \in E$ et $q + 1 \notin E$, ce qui se traduit par

$$qb \leqslant a < (q+1)b.$$

Posons $r = a - bq$. L'encadrement précédent se récrit $0 \leqslant r < b = |b|$, d'où l'existence de q et r dans ce cas.

Supposons maintenant que $b < 0$. Alors, $-b > 0$, et par le cas précédent, il existe $q, r \in \mathbb{Z}$ tels que $a = (-b)q + r = b(-q) + r$, avec $0 \leqslant r < -b = |b|$, d'où le résultat.

Montrons maintenant l'unicité de q et r. Supposons que $q, q', r, r' \in \mathbb{Z}$ vérifient

$$a = bq + r = bq' + r',$$

avec $0 \leqslant r, r' < |b|$. Alors, $-|b| < r - r' < |b|$, c'est-à-dire $|r - r'| < |b|$. Par ailleurs, $r - r' = b(q' - q)$. Si $q' - q \neq 0$, on obtient la contradiction

$$|r - r'| \geqslant |b|.$$

Ainsi, $q = q'$ et par suite $r = r'$. Cela achève la démonstration. □

2.3. Définition. Les entiers q et r définis dans le théorème précédent s'appellent respectivement le *quotient* et le *reste* de la division euclidienne de a par b.

2.4. Exemples

(1) Lorsque $a \geqslant 0$ et $b > 0$, la division euclidienne de a par b est facile à effectuer. Le quotient est le plus grand entier $q \geqslant 0$ tel que $0 \leqslant qb \leqslant a$, et le reste est alors $r = a - qb$.

Par exemple, la division euclidienne de 5 par 2 est

$$5 = 2 \times 2 + 1.$$

Autrement dit, $q = 2$ et $r = 1$.

(2) Il faut en revanche faire un peu attention lorsque l'on travaille avec des entiers négatifs. Par exemple, la division euclidienne de -5 par 2 est

$$-5 = 2\times(-3) + 1.$$

Ainsi, dans ce cas, on a $q = -3$ et $r = 1$.

Nous décrivons maintenant l'algorithme d'Euclide[1], qui démontre l'existence d'un pgcd de deux entiers (dont au moins un des deux est non nul).

Algorithme d'Euclide. Soient $a, b \in \mathbb{Z}$, avec $b \neq 0$. On définit quatre suites $(r_i)_{i\geqslant -1}$, $(q_i)_{i\geqslant 0}$, $(u_i)_{i\geqslant -1}$, $(v_i)_{i\geqslant -1}$ d'éléments de $\mathbb{Z}$ comme suit :

$$r_{-1} = a, \quad r_0 = b, \quad r_{i+1} = r_{i-1} - q_i r_i,$$

$$u_{-1} = 1, \quad u_0 = 0, \quad u_{i+1} = u_{i-1} - q_i u_i,$$

$$v_{-1} = 0, \quad v_0 = 1, \quad v_{i+1} = v_{i-1} - q_i v_i,$$

où q_i et r_{i+1} sont respectivement le quotient et le reste de la division euclidienne de r_{i-1} par r_i.

On montre que r_i est nul à partir d'un certain rang. Soit $n \geqslant 0$ le plus grand entier tel que $r_n \neq 0$, et soient

$$d = r_n, \quad u = u_n, \quad v = v_n.$$

Alors, d est un pgcd de a et b, et l'on a $ua + vb = d$.

Nous ne démontrerons pas les affirmations précédentes ici, car nous le ferons dans un contexte plus général (proposition XVIII-3.7).

On démontre également que si d est un pgcd de a et b, $\dfrac{ab}{d}$ est un ppcm de a et b.

En particulier, si $(a, b) \neq (0, 0)$, un pgcd et un ppcm de a et b existent. Un pgcd et un ppcm étant uniques au signe près (cf. remarque 1.4 (5)), il existe un unique pgcd et ppcm positifs, que l'on note respectivement $\mathrm{pgcd}(a, b)$ et $\mathrm{ppcm}(a, b)$.

2.5. Remarque. Supposons que $(a, b) \neq (0, 0)$. Alors, $\mathrm{pgcd}(a, b)$ est le plus grand diviseur positif commun à a et b pour la relation de divisibilité.

Il se trouve que c'est aussi le plus grand diviseur positif commun à a et b pour la relation d'ordre naturelle.

En effet, notons d le plus grand diviseur positif commun à a et b pour la relation d'ordre naturelle. Puisque d divise a et b, on a $d \mid \mathrm{pgcd}(a, b)$, et

1. L'algorithme présenté ici est plutôt appelé l'algorithme d'Euclide **étendu**, l'algorithme d'Euclide classique consistant seulement à calculer la suite des restes $(r_i)_{i\geqslant -1}$.

ainsi $d \leqslant \text{pgcd}(a,b)$ d'après la remarque 1.2 (4). Or, par définition de d, on a aussi $\text{pgcd}(a,b) \leqslant d$, d'où l'égalité désirée.

Remarquons néanmoins que la définition d'un pgcd en termes de divisibilité est « la bonne », dans le sens où la relation de divisibilité se généralise de façon raisonnable dans n'importe quel anneau, contrairement à la notion d'élément positif.

3. Entiers premiers entre eux

3.1. Définition. Soient a, $b \in \mathbb{Z}$. On dit que a et b sont *premiers entre eux* si leurs seuls diviseurs communs sont 1 et -1.

Cela implique en particulier que $(a,b) \neq (0,0)$. On peut donc reformuler la définition comme suit : a et b sont premiers entre eux si $\text{pgcd}(a,b) = 1$.

Remarquons que l'algorithme d'Euclide fournit en partie le résultat suivant.

3.2. Théorème (Bézout). *Soient a, $b \in \mathbb{Z}$, $(a,b) \neq (0,0)$.*

(1) *Soit d un pgcd de a et b. Alors, il existe $u, v \in \mathbb{Z}$ tels que $ua + vb = d$ (Relation de Bézout).*

(2) *Les entiers a et b sont premiers entre eux si, et seulement si, il existe des entiers u, $v \in \mathbb{Z}$ tels que $ua + vb = 1$.*

Démonstration. Le point (1) découle directement de l'algorithme d'Euclide. Si maintenant a et b sont des entiers premiers entre eux, les pgcd de a et b sont 1 et -1. On applique alors le point (1) pour obtenir le sens direct du point (2), quitte à remplacer (u,v) par $(-u,-v)$. Réciproquement, supposons qu'il existe $u, v \in \mathbb{Z}$ tels que $ua + vb = 1$. Soit $d \in \mathbb{Z}$ un diviseur commun à a et b. Puisque $d \mid a$ et $d \mid b$, on a $d \mid ua + vb$, c'est-à-dire $d \mid 1$. Comme on a toujours $1 \mid d$, on obtient $d = \pm 1$, par la remarque 1.2 (1). Ainsi, a et b sont premiers entre eux, d'où le résultat. □

On finit ce paragraphe par un lemme dû à Gauss.

3.3. Lemme (Gauss). *Soient a, $b \in \mathbb{Z}$ deux entiers premiers entre eux. Alors, pour tout $c \in \mathbb{Z}$, $a \mid bc \implies a \mid c$.*

Démonstration. Par le théorème de Bézout, il existe deux entiers u, $v \in \mathbb{Z}$ tels que $ua + vb = 1$. On a alors $uac + vbc = c$. Comme $a \mid ac$ et $a \mid bc$, on obtient $a \mid uac + vbc$, c'est-à-dire $a \mid c$. □

4. Décomposition en facteurs premiers

On définit maintenant la notion de nombre premier.

4.1. Définition. Un nombre $p \in \mathbb{N}$ est dit *premier* si $p \geqslant 2$ et si ses seuls diviseurs positifs sont 1 et p.

Soit $a \in \mathbb{Z}$. Alors, $\operatorname{pgcd}(a,p)$ est en particulier un diviseur positif de p, et vaut donc 1 ou p. Comme $\operatorname{pgcd}(a,p)$ est aussi un diviseur de a, on en déduit immédiatement le résultat suivant.

4.2. Lemme. *Soit p un nombre premier. Alors, pour tout $a \in \mathbb{Z}$, $p \nmid a$ si, et seulement si, p et a sont premiers entre eux.*

On peut alors en déduire une propriété fondamentale des nombres premiers.

4.3. Lemme (Euclide). *Soit p un nombre premier.*
Alors, pour tous $a, b \in \mathbb{Z}$, on a

$$p \mid ab \quad \Longrightarrow \quad (p \mid a \ \text{ ou } \ p \mid b).$$

Démonstration. Soient $a, b \in \mathbb{Z}$ tels que $p \mid ab$. Si $p \mid a$, il n'y a rien à faire. Si $p \nmid a$, le lemme précédent montre que a et p sont premiers entre eux. Mais alors, $p \mid b$, par le lemme de Gauss. □

On en vient maintenant au résultat principal de ce paragraphe.

4.4. Théorème (Théorème fondamental de l'arithmétique). *Tout entier $a \in \mathbb{Z}$ non nul s'écrit de manière unique sous la forme*

$$a \ = \ \varepsilon \prod_p p^{n_p},$$

où p décrit l'ensemble des nombres premiers, $\varepsilon = \pm 1$, et les entiers $n_p \geqslant 0$ sont tous nuls sauf un nombre fini.

Démonstration. Nous esquissons seulement la démonstration, car nous obtiendrons ce résultat comme cas particulier d'un théorème plus général [2].

L'existence se démontre par récurrence. Quitte à remplacer a par $-a$, on peut supposer $a \geqslant 1$. Pour $n \geqslant 1$, soit (H_n) la proposition suivante.

(H_n) Tout entier $a \in [\![1, n]\!]$ s'écrit comme un produit fini de nombres premiers (éventuellement vide).

2. On montrera en fait que tout anneau ayant une division euclidienne jouit d'une propriété similaire (proposition XVIII-3.3).

Alors, (H_1) est vraie. Supposons que (H_{n-1}) soit vraie pour un entier $n \geqslant 2$. Montrons que (H_n) est vraie. Il suffit de vérifier que n s'écrit comme un produit fini de nombres premiers. Soit p le plus petit entier[3] strictement supérieur à 1 divisant n. Alors, p est un nombre premier. En effet, tout entier naturel divisant p divise n, donc vaut 1 ou p par minimalité de p. Comme $\frac{n}{p} \leqslant \frac{n}{2} < n$, l'hypothèse de récurrence implique que $\frac{n}{p}$ est un produit de nombres premiers. Il en est donc de même de $n = p\frac{n}{p}$. Par conséquent, (H_n) est vraie, et cela achève la récurrence.

On obtient alors la décomposition annoncée en regroupant les termes. L'unicité s'obtient par une utilisation répétée du lemme d'Euclide. □

4.5. Définition. Soit $a \in \mathbb{Z}$ un entier non nul, et soit

$$a = \varepsilon \prod_p p^{n_p}$$

sa décomposition en facteurs premiers. Pour tout p premier, l'entier n_p est appelé la *valuation p-adique de a*, et est noté $v_p(a)$.

On pose aussi $v_p(0) = +\infty$ pour tout p premier.

La démonstration du lemme suivant est laissée au lecteur.

4.6. Lemme. *Soit p un nombre premier. Alors :*

(1) *pour tout $a \in \mathbb{Z}$, on a $v_p(a) = +\infty$ si, et seulement si, $a = 0$;*

(2) *pour tous $a, b \in \mathbb{Z}$, on a $v_p(a+b) \geqslant \min\big(v_p(a), v_p(b)\big)$, avec égalité dès que $v_p(a) \neq v_p(b)$;*

(3) *pour tous $a, b \in \mathbb{Z}$, on a $v_p(ab) = v_p(a) + v_p(b)$;*

De plus, pour tous $a, b \in \mathbb{Z}$, on a $a \mid b$ si, et seulement si, $v_p(a) \leqslant v_p(b)$ pour tout p premier.

La démonstration du corollaire suivant est aussi laissée en exercice au lecteur.

4.7. Corollaire. *Soient $a, b \in \mathbb{Z}$. Alors, on a*

$$\operatorname{pgcd}(a,b) = \prod_p p^{\min(v_p(a),v_p(b))} \quad \text{et} \quad \operatorname{ppcm}(a,b) = \prod_p p^{\max(v_p(a),v_p(b))}.$$

3. L'ensemble des diviseurs > 1 de n est en effet non vide, puisque $n \mid n$.

5. Les pgcd et ppcm d'une famille quelconque d'entiers

Nous allons maintenant généraliser la notion de pgcd et ppcm à une famille quelconque d'entiers, éventuellement infinie.

5.1. Définition. Soit $(a_i)_{i\in I}$ une famille d'entiers. On dit que $d \in \mathbb{Z}$ est un *pgcd* de la famille d'entiers $(a_i)_{i\in I}$ si d vérifie les propriétés suivantes :

(1) on a $d \mid a_i$ pour tout $i \in I$;

(2) pour tout $c \in \mathbb{Z}$ tel que $c \mid a_i \in I$ pour tout $i \in I$, on a $c \mid d$.

On dit que $m \in \mathbb{Z}$ est un *ppcm* de la famille d'entiers $(a_i)_{i\in I}$ si m vérifie les propriétés suivantes :

(1) on a $a_i \mid m$ pour tout $i \in I$;

(2) pour tout $c \in \mathbb{Z}$ tel que $a_i \mid c$ pour tout $i \in I$, on a $m \mid c$.

Un pgcd ou un ppcm, lorsqu'il existe, est unique au signe près. On écrit alors $\operatorname{pgcd}\big((a_i)_{i\in I}\big)$ et $\operatorname{ppcm}\big((a_i)_{i\in I}\big)$ pour l'unique pgcd et ppcm positifs.

Le résultat suivant généralise le corollaire 4.7, et est laissé en exercice au lecteur.

5.2. Lemme. *Soit $(a_i)_{i\in I}$ une famille d'entiers non nuls. Écrivons*

$$a_i = \varepsilon_i \prod_p p^{v_p(a_i)},$$

pour tout $i \in I$. Alors, on a

$$\operatorname{pgcd}\big((a_i)_{i\in I}\big) = \prod_p p^{\min_{i\in I}(v_p(a_i))}.$$

De plus, un ppcm de la famille $(a_i)_{i\in I}$ existe si, et seulement si, $(a_i)_{i\in I}$ est bornée. Dans ce cas, on a

$$\operatorname{ppcm}\big((a_i)_{i\in I}\big) = \prod_p p^{\max_{i\in I}(v_p(a_i))}.$$

5.3. Corollaire. *Soit $(a_i)_{i\in I}$ une famille d'entiers non nuls. Soit $I = \bigcup_{j\in J} I_j$ une partition de I. Alors, on a*

$$\operatorname{pgcd}\big((a_i)_{i\in I}\big) = \operatorname{pgcd}\Big(\big(\operatorname{pgcd}\big((a_i)_{i\in I}\big)\big)_{j\in J}\Big).$$

Si de plus $(a_i)_{i\in I}$ est bornée, on a

$$\mathrm{ppcm}\big((a_i)_{i\in I}\big) = \mathrm{ppcm}\Big(\big(\mathrm{ppcm}\big((a_i)_{i\in I}\big)\big)_{j\in J}\Big).$$

Démonstration. Cela provient des propriétés du maximum et du minimum d'une famille d'entiers, et du corollaire précédent. □

5.4. Remarque. Lorsque la famille $(a_i)_{i\in I}$ est finie, le résultat précédent montre que l'on peut ramener le calcul du pgcd à des calculs successifs de pgcd de deux entiers. En particulier, on peut le calculer en utilisant l'algorithme d'Euclide.

6. Arithmétique modulaire

On termine ce chapitre en donnant quelques compléments sur la relation de congruence modulo un entier.

6.1. Définition. Soit $n \geqslant 0$ un entier. On dit que $a, b \in \mathbb{Z}$ sont *congrus modulo n* si $a - b \in \mathbb{Z}$ est un multiple de n, i.e. s'il existe un entier $k \in \mathbb{Z}$ tel que $a = b + kn$.

On le note $a \equiv b\ [n]$.

Le lemme suivant résume les propriétés calculatoires de la relation de congruence modulo n.

6.2. Lemme. *Soit $n \geqslant 0$. La relation de congruence modulo n est une relation d'équivalence sur $\mathbb{Z}$.*

De plus, pour tous $a, a', b, b' \in \mathbb{Z}$ tels que $a \equiv a'\ [n]$ et $b \equiv b'\ [n]$, on a

(1) $a + b \equiv a' + b'\ [n]$;

(2) $-a \equiv -a'\ [n]$;

(3) $ab \equiv a'b'\ [n]$.

Démonstration. Montrons que la relation $\equiv$ est une relation d'équivalence. Soient a, b, $c \in \mathbb{Z}$. Puisque $a - a = 0$ est un multiple de n, on a bien $a \equiv a\ [n]$.

Supposons maintenant que $a \equiv b\ [n]$. Alors, $a - b$ est un multiple de n, et donc $b - a = -(a - b)$ aussi. Ainsi, $b \equiv a\ [n]$.

Supposons enfin que $a \equiv b\ [n]$ et $b \equiv c\ [n]$. Les entiers $a - b$ et $b - c$ sont donc des multiples de n. Mais alors, $a - c = (a - b) + (b - c)$ est encore un multiple de n, d'où $a \equiv c\ [n]$.

Ainsi, $\equiv$ est bien une relation d'équivalence. Supposons maintenant que l'on ait $a \equiv a'\ [n]$ et $b \equiv b'\ [n]$. Montrons les propriétés (1) – (3).

Par hypothèse, il existe des entiers $r,\ s \in \mathbb{Z}$ tels que $a = a' + rn$ et $b = b' + sn$. On a alors

$$a+b=a'+b'+(r+s)n,\ -a=-a'+(-r)n \quad \text{et} \quad ab=a'b'+(sa'+rb'+rsn)n.$$

Comme $r+s,\ -r$ et $sa'+rb'+rsn$ sont des entiers, on en déduit le résultat voulu. □

D'après le lemme précédent, la relation de congruence modulo n est une relation d'équivalence sur $\mathbb{Z}$. On note $\mathbb{Z}/n\mathbb{Z}$ l'ensemble quotient. Si $a \in \mathbb{Z}$, on note $\overline{a}$ la classe d'équivalence correspondante.

Nous allons maintenant montrer que l'on peut munir $\mathbb{Z}/n\mathbb{Z}$ d'une structure d'anneau. Comme les éléments de $\mathbb{Z}/n\mathbb{Z}$ sont des classes d'équivalence d'éléments de $\mathbb{Z}$, l'idée la plus naturelle est d'exploiter la structure d'anneau de $\mathbb{Z}$. Si ω et ω' sont deux éléments de $\mathbb{Z}/n\mathbb{Z}$, il existe par définition des entiers $a,\ b \in \mathbb{Z}$ tels que $\omega = \overline{a}$ et $\omega' = \overline{b}$. Il semble alors sensé de poser

$$\omega + \omega' = \overline{a+b} \quad \text{et} \quad \omega \cdot \omega' = \overline{ab}.$$

La difficulté est qu'une même classe possède a priori plusieurs représentants (ici, il y en a même une infinité). Il faut donc démontrer que le résultat ne dépend pas des représentants choisis.

Dans le cas présent, on doit donc montrer que si a' et b' sont d'autres représentants de ω et ω', i.e. si $\omega = \overline{a} = \overline{a'}$ et $\omega' = \overline{b} = \overline{b'}$, on a

$$\overline{a+b} = \overline{a'+b'} \quad \text{et} \quad \overline{ab} = \overline{a'b'}.$$

Or, par définition d'une classe d'équivalence, on a $\overline{a} = \overline{a'}$ si, et seulement si, $a \equiv a'\ [n]$. De même, on a $b \equiv b'\ [n]$. Par le lemme précédent, on a

$$a + b \equiv a' + b'\ [n] \quad \text{et} \quad ab \equiv a'b'\ [n].$$

Cela revient à dire que $\overline{a+b} = \overline{a'+b'}$ et $\overline{ab} = \overline{a'b'}$, d'où le résultat.

En résumé, en utilisant le fait que tout élément de l'ensemble $\mathbb{Z}/n\mathbb{Z}$ est de la forme $\overline{a}, a \in \mathbb{Z}$, on voit que l'on peut définir la somme et le produit de deux classes par

$$\overline{a} + \overline{b} = \overline{a+b} \quad \text{et} \quad \overline{a} \cdot \overline{b} = \overline{ab}.$$

Attention ! Cette formulation plus simple ne doit pas cacher la difficulté exposée plus haut. L'élément $\overline{a}$ est une classe d'équivalence, et malgré la notation trompeuse, a n'est pas défini de manière unique. Il a donc fallu montrer que la somme et le produit de deux classes étaient bien définis en vérifiant que les résultats ne dépendaient pas du choix des représentants.

L'ensemble $\mathbb{Z}/n\mathbb{Z}$, muni de ces deux lois, a les propriétés attendues.

6.3. Proposition. *Soit $n \geqslant 0$. Les lois*

$$\begin{aligned} \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} &\longrightarrow \mathbb{Z}/n\mathbb{Z} \\ (\overline{a}, \overline{b}) &\longmapsto \overline{a} + \overline{b} = \overline{a+b} \end{aligned}$$

et

$$\begin{aligned} \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} &\longrightarrow \mathbb{Z}/n\mathbb{Z} \\ (\overline{a}, \overline{b}) &\longmapsto \overline{a} \cdot \overline{b} = \overline{ab} \end{aligned}$$

confèrent à $\mathbb{Z}/n\mathbb{Z}$ une structure d'anneau, dont les éléments neutres pour l'addition et le produit sont respectivement $\overline{0}$ et $\overline{1}$, et l'opposé de $\overline{a}$ est $\overline{-a}$.

Démonstration. Toutes les propriétés que l'on doit vérifier découlent en fait des définitions et des propriétés de $\mathbb{Z}$.

Montrons par exemple l'associativité de l'addition. Pour tous a, b, $c \in \mathbb{Z}$, il vient :

$$\begin{aligned} (\overline{a} + \overline{b}) + \overline{c} &= \overline{a+b} + \overline{c} &&= \overline{(a+b)+c} \\ &= \overline{a + (b+c)} &&= \overline{a} + \overline{b+c} \\ &= \overline{a} + (\overline{b} + \overline{c}). \end{aligned}$$

On démontre de même toutes les autres propriétés requises. □

6.4. Remarque. Pour tout $a \in \mathbb{Z}$, on a l'équivalence : $\overline{a} = \overline{0} \iff n \mid a$.

Cela provient en effet de la définition de $\mathbb{Z}/n\mathbb{Z}$. Il est néanmoins **fondamental** de garder cette propriété en tête pour raisonner dans cet anneau. Nous utiliserons souvent cette remarque sans y faire référence explicitement.

Nous finissons ce chapitre en donnant quelques propriétés élémentaires de l'anneau $\mathbb{Z}/n\mathbb{Z}$. Une étude plus approfondie de cet anneau sera proposée dans le chapitre XVII.

6.5. Lemme. *Pour tout $n \geqslant 1$, l'anneau $\mathbb{Z}/n\mathbb{Z}$ possède n éléments, qui sont $\overline{0}, \overline{1}, \ldots, \overline{n-1}$.*

Démonstration. Commençons par montrer que les classes $\overline{0}, \overline{1}, \ldots, \overline{n-1}$ sont distinctes. Soient $i, j \in [\![0, n-1]\!]$ tels que $\overline{i} = \overline{j}$. Alors, $i \equiv j\ [n]$, et donc $i - j$ est un multiple de n. Mais, on a

$$-(n-1) \leqslant i - j \leqslant n - 1,$$

et donc nécessairement $i - j = 0$, soit $i = j$, d'où le résultat.

Il reste à vérifier que ces n classes sont les seuls éléments de l'anneau $\mathbb{Z}/n\mathbb{Z}$. Soit $\overline{a} \in \mathbb{Z}/n\mathbb{Z}$. Écrivons la division euclidienne de a par n. Il existe $q, r \in \mathbb{Z}$, avec $0 \leqslant r < n$ tels que $a = nq + r$. Mais alors, $a \equiv r\ [n]$, et donc $\overline{a} = \overline{r}$. Comme $r \in [\![0, n-1]\!]$, cela achève la démonstration. □

6.6. Remarque. Ce résultat est faux si $n = 0$. En effet, pour tous a, $b \in \mathbb{Z}$, on a $a \equiv b\ [0] \iff a = b$. Autrement dit, $\overline{a} = \{a\}$, et $\mathbb{Z}/0\mathbb{Z}$ possède une infinité d'éléments.
Cet anneau n'est d'ailleurs rien d'autre que l'anneau $\mathbb{Z}$, de manière un peu déguisée.

On va pouvoir maintenant donner l'exemple de toute une famille de corps un peu « exotiques ».

6.7. Proposition. *Soit p un nombre premier. Alors, l'anneau $\mathbb{Z}/p\mathbb{Z}$ est un corps.*

Démonstration. Soit p un nombre premier. En particulier, $p \geqslant 2$, d'où $\overline{0} \neq \overline{1}$ par le lemme 6.5. Soit $\overline{a} \in \mathbb{Z}/p\mathbb{Z}$ non nul. Puisque $\overline{a} \neq \overline{0}$, a n'est pas un multiple de p. Par le lemme 4.2, a et p sont premiers entre eux. Par le théorème de Bézout, il existe donc u, $v \in \mathbb{Z}$ tels que

$$ua + vp = 1.$$

Cela revient à dire que $ua \equiv 1\ [p]$, ou encore $\overline{ua} = \overline{1}$. Mais alors,

$$\overline{a}\cdot\overline{u} = \overline{u}\cdot\overline{a} = \overline{ua} = \overline{1}.$$

Ainsi, $\overline{a}$ possède un symétrique pour la loi produit, à savoir $\overline{u}$. Cela démontre que $\mathbb{Z}/p\mathbb{Z}$ est bien un corps. □

Notation. Si p est un nombre premier, on note $\mathbb{F}_p$ le corps $\mathbb{Z}/p\mathbb{Z}$.

7. Exercices

E1. Soient a, $b \in \mathbb{Z}$, avec b non nul.

a. Quels sont le quotient et le reste de la division euclidienne de a par 1 ?

b. On note q et r le quotient et le reste de la division euclidienne de a par b. Quels sont le quotient et le reste de la division euclidienne de a par $-b$? Et de $-a$ par b (distinguer les cas $r = 0$, $r > 0$ et $b > 0$, $r > 0$ et $b < 0$) ?

c. Quel est le reste de la division euclidienne de 5 par 18 ? et de la division de -5 par -18 ?

E2. Appliquer l'algorithme d'Euclide pour trouver le pgcd de 1092 et 1732, ainsi qu'un couple de Bézout associé.

E3. Soient s, m, $n \in \mathbb{Z}$, $s \neq \pm 1$, m, $n \geqslant 1$. Soient q et r le quotient et le reste de la division euclidienne de m par n. Écrire le quotient et le reste de la division euclidienne de $s^m - 1$ par $s^n - 1$ en fonction de s, q

et r. En déduire le pgcd de $s^m - 1$ et $s^n - 1$ en fonction du pgcd de m et n.

Indication. On pourra commencer par regarder des exemples pour deviner le résultat.

E4. Soient $a, b \in \mathbb{Z}$, $(a, b) \neq (0, 0)$.

a. On suppose que a et b sont premiers entre eux. Soit $m \in \mathbb{Z}$. Utiliser le lemme de Gauss pour montrer que m est multiple de a et b si, et seulement si, m est multiple de ab.

b. On suppose que $\text{pgcd}(a, b) = d$.

(i) Montrer que si $m \in \mathbb{Z}$ est un multiple de a et b, alors $\frac{m}{d}$ est un multiple de $\frac{a}{d}$ et $\frac{b}{d}$.

(ii) En utilisant **a.**, en déduire que m est multiple de a et b si, et seulement si, m est multiple de $\frac{ab}{d}$.

(iii) Calculer $\text{ppcm}(a, b)$ en fonction de a, b et $\text{pgcd}(a, b)$.

E5. Soient $a, b, d \in \mathbb{Z}$, avec $d \geqslant 0$.

Si l'on a $ua + vb = d$ avec $u, v \in \mathbb{Z}$, a-t-on $d = \text{pgcd}(a, b)$? Et si d est un diviseur commun à a et b ?

E6. Soient $a, b \in \mathbb{Z}$ non nuls.

On suppose que, pour tout $c \in \mathbb{Z}$, $a \mid bc \Longrightarrow a \mid c$. Montrer que a et b sont premiers entre eux.

E7. Soient $a, b \in \mathbb{Z}$, $(a, b) \neq (0, 0)$. Soit (u_0, v_0) un couple de Bézout associé à a et b. Décrire tous les couples de Bézout associés à a et b en fonction de u_0 et v_0.

E8. Soient $a, b, c \in \mathbb{Z}$ non nuls. Sans utiliser la décomposition en facteurs premiers, montrer :

(1) $\text{pgcd}(a, b) = 1$ et $\text{pgcd}(a, c) = 1 \iff \text{pgcd}(a, bc) = 1$;

(2) $\text{pgcd}(ac, bc) = \text{pgcd}(a, b)|c|$;

(3) $\text{pgcd}(a, b) = 1 \iff \text{pgcd}(a^n, b^m) = 1$, pour tous $m, n \geqslant 1$.

E9. Démontrer le corollaire 4.7.

E10. **a.** Montrer que tout rationnel $r \in \mathbb{Q}$ non nul s'écrit de manière unique sous la forme

$$r = \varepsilon \prod_p p^{n_p},$$

où $\varepsilon \in \{\pm 1\}$, $n_p \in \mathbb{Z}$, et les entiers n_p sont presque tous nuls.

L'entier $n_p \in \mathbb{Z}$ est encore noté $v_p(r)$. On pose aussi $v_p(0) = +\infty$.

b. Soit p un nombre premier. Montrer les propriétés suivantes :

(i) pour tout $r \in \mathbb{Q}$, on a $v_p(r) = +\infty$ si, et seulement si, $r = 0$;

(ii) pour tous r, $s \in \mathbb{Q}$, on a $v_p(r+s) \geqslant \min(v_p(r), v_p(s))$, avec égalité dès que $v_p(r) \neq v_p(s)$;

(iii) pour tous r, $s \in \mathbb{Q}$, on a $v_p(rs) = v_p(r) + v_p(s)$.

(iv) pour tout $r \in \mathbb{Q}$, on a $r \in \mathbb{Z}$ si, et seulement si, $v_p(r) \geqslant 0$ pour tout p.

(v) pour tous $a, b \in \mathbb{Z}$, on a $a \mid b$ si, et seulement si, $v_p(b) \geqslant v_p(a)$ pour tout p.

E11. Sans utiliser la décomposition en facteurs premiers, montrer que $\sqrt{2}$ est irrationnel.

E12. Justifier le critère de divisibilité par 4 et la preuve par 9.

E13. Montrer qu'il existe une infinité de nombres premiers.

Indication. Si $p_1, \ldots, p_k$ sont des nombres premiers distincts, on pourra considérer un diviseur premier de $1 + p_1 \cdots p_k$.

E14.

a. En calculant dans $\mathbb{Z}/17\mathbb{Z}$, calculer le reste de la division euclidienne de $16 \cdot 15 \cdot 14$ par 17.

b. Calculer $\overline{2}^m$ pour $m \in [\![1, 8]\!]$. Déterminer le reste de la division euclidienne de 2^{2014} par 17.

c. Le couple $(\mathbb{Z}/n\mathbb{Z} \setminus \{\overline{0}\}, \cdot)$ est-il un groupe ?

d. Vrai ou faux : si $\overline{x} \cdot \overline{y} = \overline{0}$ dans $\mathbb{Z}/n\mathbb{Z}$, on a $\overline{x} = \overline{0}$ ou $\overline{y} = \overline{0}$.

e. Soit $m \in \mathbb{Z}$. Trouver une condition nécessaire et suffisante sur m pour que $\overline{m} \in \mathbb{Z}/n\mathbb{Z}$ soit inversible pour la multiplication.

E15. Soient a_1, a_2, m_1, $m_2 \in \mathbb{Z}$, avec m_1, $m_2 \geqslant 1$. On considère le système de deux congruences

$$\begin{cases} x \equiv a_1 \ [m_1] \\ x \equiv a_2 \ [m_2] \end{cases}, \quad x \in \mathbb{Z}.$$

a. On suppose que m_1 et m_2 sont premiers entre eux. Montrer que le système admet toujours une solution (on pourra commencer par les cas $(a_1, a_2) = (1, 0)$ et $(a_1, a_2) = (0, 1)$). Décrire alors l'ensemble des solutions.

b. On suppose que $\operatorname{pgcd}(m_1, m_2) = d$. Montrer que si le système précédent a une solution, alors $a_1 \equiv a_2 \mod d$. Montrer que cette condition est aussi suffisante (on pourra poser $y = x - a_1$). Décrire alors l'ensemble des solutions.

E16. Deux joueurs jouent au jeu suivant : on dispose de vingt allumettes alignées sur une table. Chaque joueur enlève à tour de rôle une, deux ou trois allumettes. Le joueur qui prend la dernière allumette a perdu. Montrer que le premier joueur dispose d'une stratégie gagnante.

E17. Sur une île vit une population de caméléons : treize rouges, quinze bleus et dix-sept verts. Chaque fois que deux caméléons de couleurs différentes se croisent, ils prennent tous les deux la troisième couleur. Est-il possible qu'au bout d'un certain temps, les caméléons de l'île soient tous de la même couleur ?

Indication. Soit r le nombre de caméléons rouges et soit b le nombre de caméléons bleus. Comment varie la quantité $r - b$ lorsque deux caméléons de deux couleurs différentes se croisent ?

E18. Cent prisonniers sont convoqués un soir par le directeur de la prison. Celui-ci leur dit : « Demain, vous serez alignés à la queue-leu-leu avec un chapeau de couleur sur la tête (jaune, bleu ou vert), sans pouvoir vous retourner, et vous devrez deviner la couleur de votre chapeau. Bien sûr, vous n'aurez droit qu'à une seule réponse. Tous ceux qui réussiront seront graciés, les autres seront exécutés. »

Les prisonniers se réunissent et établissent une stratégie qui leur permet de sauver quatre-vingt-dix-neuf d'entre eux. Laquelle ? Généralisation avec n prisonniers et q couleurs ? On suppose que chaque prisonnier peut voir la couleur des chapeaux des prisonniers se trouvant devant lui.

Chapitre III

Rappels et compléments d'algèbre linéaire

1. Généralités

Commençons par rappeler la définition d'un espace vectoriel sur un corps.

1.1. Définition. Soit K un corps. Un *K-espace vectoriel* est un ensemble E, muni de deux opérations

$$\begin{array}{ccc} E \times E \longrightarrow E & & K \times E \longrightarrow E \\ (x,y) \longmapsto x+y & \text{et} & (\lambda, x) \longmapsto \lambda \cdot x \end{array}$$

vérifiant les propriétés suivantes :

(1) le couple $(E,+)$ est un groupe abélien. Son élément neutre est alors noté 0_E, ou même 0, et le symétrique de $x \in E$ est noté $-x$, et est appelé *opposé* de x ;
(2) on a $1 \cdot x = x$, pour tout $x \in E$;
(3) on a $(\lambda\mu)\cdot x = \lambda \cdot (\mu \cdot x)$, pour tous λ, $\mu \in K$, et tout $x \in E$;
(4) on a $\lambda \cdot (x+y) = \lambda \cdot x + \lambda \cdot y$, pour tout $\lambda \in K$, et tous x, $y \in E$;
(5) on a $(\lambda+\mu)\cdot x = \lambda \cdot x + \mu \cdot x$, pour tous λ, $\mu \in K$, et tout $x \in E$.

En particulier, on a $0 \cdot x = 0$ pour tout $x \in E$.

Un élément de E est appelé un *vecteur*, et 0 est appelé le *vecteur nul*.

1.2. Exemple. L'ensemble $\mathscr{F}(X,K)$ des fonctions d'un ensemble X vers K est un K-espace vectoriel, de même que K^n, pour tout $n \geqslant 1$.

1.3. Définition. Soit E un K-espace vectoriel. Un *sous-espace vectoriel* est un sous-ensemble $F \subset E$ vérifiant :

(1) on a $0 \in F$;
(2) pour tous $x, y \in F$, on a $x + y \in F$;
(3) pour tout $\lambda \in K$, et pour tout $x \in F$, on a $\lambda \cdot x \in F$.

De manière équivalente, un sous-espace vectoriel est un sous-ensemble F non vide tel que pour tous $x, y \in F$ et tout $\lambda \in K$, on a $x + \lambda \cdot y \in F$. Un sous-espace vectoriel de E est alors un K-espace vectoriel pour les lois induites sur F par celles de E.

On peut montrer aisément que l'intersection d'une famille de sous-espaces vectoriels de E est encore un sous-espace vectoriel de E. On a alors la définition suivante.

1.4. Définition. Soit E un K-espace vectoriel, et soit S une partie de E. Le *sous-espace vectoriel engendré par* S est l'intersection de tous les sous-espaces vectoriels de E contenant S ; c'est le plus petit sous-espace vectoriel de E contenant S. Il est noté $\mathrm{Vect}_K(S)$.

On a alors une description explicite du sous-espace vectoriel engendré par une partie.

1.5. Proposition. *Soit E un K-espace vectoriel, et soit S une partie (éventuellement vide) de E. Alors, on a*

$$\mathrm{Vect}_K(S) = \Big\{ \sum_{i=1}^{n} \lambda_i \cdot x_i \mid n \geqslant 0,\ x_i \in S,\ \lambda_i \in K \Big\}.$$

Passons maintenant aux applications linéaires.

1.6. Définition. Soient E, F deux K-espaces vectoriels. Une *application linéaire* est une application $u : E \longrightarrow F$ telle que

(1) pour tous $x, y \in E$, on a $u(x + y) = u(x) + u(y)$;
(2) pour tout $\lambda \in K$, et tout $x \in E$, on a $u(\lambda \cdot x) = \lambda \cdot u(x)$.

De manière équivalente, $u : E \longrightarrow F$ est linéaire si pour tous $x, y \in E$, et tout $\lambda \in K$, on a

$$u(x + \lambda \cdot y) = u(x) + \lambda \cdot u(y).$$

Un *endomorphisme* de E est une application linéaire de E dans lui-même.

L'ensemble des applications linéaires de E vers F est un K-espace vectoriel pour les lois évidentes, noté $\mathscr{L}(E, F)$. Si $E = F$, on le note plus simplement $\mathscr{L}(E)$. L'ensemble $\mathscr{L}(E)$, muni de l'addition et de la composition des

endomorphismes, est un anneau (cela utilise de manière cruciale la linéarité d'un endomorphisme).

Un *isomorphisme* d'espaces vectoriels est une application linéaire bijective. Dans ce cas, on peut montrer que l'application inverse est aussi linéaire. Un *automorphisme* de E est un isomorphisme de E sur lui-même. L'ensemble des automorphismes de E est un groupe pour la composition des applications, noté $\mathrm{GL}(E)$, et appelé *groupe linéaire.*

On rappelle maintenant la notion de K-algèbre.

1.7. Définition. Soit K un corps. Une *K-algèbre* est un K-espace vectoriel $\mathscr{A}$, muni d'une application K-bilinéaire

$$\begin{aligned} \mathscr{A} \times \mathscr{A} &\longrightarrow \mathscr{A} \\ (x, y) &\longmapsto xy, \end{aligned}$$

appelée *loi produit.* On dit que $\mathscr{A}$ est *associative* si le produit est associatif, et *unitaire* si le produit admet un élément neutre, que l'on note $1_{\mathscr{A}}$.

Un *morphisme* de K-algèbres $f : \mathscr{A} \longrightarrow \mathscr{B}$ est une application K-linéaire vérifiant

$$f(xy) = f(x)f(y), \quad \text{pour tous } x, y \in \mathscr{A}.$$

Si $\mathscr{A}$ et $\mathscr{B}$ sont unitaires, on impose la condition supplémentaire

$$f(1_{\mathscr{A}}) = 1_{\mathscr{B}}.$$

On laisse le soin au lecteur de définir la notion d'endomorphisme, d'isomorphisme ou d'automorphisme de K-algèbres.

Enfin, une *sous-algèbre* de $\mathscr{A}$ est un sous-espace vectoriel $\mathscr{B}$ de $\mathscr{A}$ qui est stable par produit. Si $\mathscr{A}$ est unitaire, on demande aussi que $1_{\mathscr{A}} \in \mathscr{B}$.

1.8. Exemples

(1) Pour tout $n \geqslant 1$, $\mathrm{M}_n(K)$ est une algèbre associative unitaire pour le produit matriciel.

(2) Plus généralement, si E est un K-espace vectoriel, alors le K-espace vectoriel $\mathscr{L}(E)$ est une K-algèbre associative unitaire pour la composition.

(3) Pour tout ensemble non vide Ω, $\mathscr{F}(\Omega, K)$ est une K-algèbre associative unitaire pour le produit des fonctions.

(4) L'anneau des polynômes $K[X]$ est une K-algèbre associative unitaire pour le produit des polynômes.

(5) Le $\mathbb{R}$-espace vectoriel $\mathbb{R}^3$, muni du produit vectoriel, est une $\mathbb{R}$-algèbre non associative, et non unitaire.

(6) Une sous-algèbre d'une K-algèbre (associative unitaire) $\mathscr{A}$ est aussi une K-algèbre (associative unitaire) pour les lois induites par celles de la K-algèbre $\mathscr{A}$.

1.9. Remarque. Si $\mathscr{A}$ est une K-algèbre associative unitaire, alors le groupe additif $\mathscr{A}$ est un anneau pour la loi produit, et l'application

$$\begin{aligned} K &\longrightarrow \mathscr{A} \\ \lambda &\longmapsto \lambda \cdot 1_{\mathscr{A}} \end{aligned}$$

est un morphisme d'anneaux, dont l'image commute avec les éléments de $\mathscr{A}$.

Inversement, si A est un anneau, muni d'un morphisme d'anneaux

$$\iota : K \longrightarrow A$$

dont l'image commute avec les éléments de A, alors A peut être canoniquement muni d'une structure de K-algèbre associative unitaire. En effet, la loi externe

$$\begin{aligned} K \times A &\longrightarrow A \\ (\lambda, x) &\longmapsto \lambda \cdot x = \iota(\lambda) x \end{aligned}$$

muni $(A, +)$ d'une structure de K-espace vectoriel, pour laquelle le produit dans A est K-bilinéaire.

Les deux constructions sont inverses l'une de l'autre.

Dans la suite, s'il n'y a pas d'ambiguïté, nous écrirons « λx » à la place de « $\lambda \cdot x$ », lorsque $\lambda \in K$ et x est un élément d'un K-espace vectoriel.

2. Familles libres, génératrices, bases

Nous allons maintenant donner quelques compléments sur la notion de famille libre, génératrice, et de base.

2.1. Définition. Soit E un K-espace vectoriel. Une famille $(e_i)_{i \in I}$ d'éléments de E est *libre* si pour tout sous-ensemble $J \subset I$ **fini**, et pour toute famille $(\lambda_j)_{j \in J}$ d'éléments de K, on a

$$\sum_{j \in J} \lambda_j e_j = 0 \quad \Longrightarrow \quad \lambda_j = 0, \ \text{ pour tout } j \in J.$$

De manière équivalente, $(e_i)_{i \in I}$ est libre si pour toute famille $(\lambda_i)_{i \in I}$ d'éléments de K presque tous nuls, on a

$$\sum_{i \in I} \lambda_i e_i = 0 \quad \Longrightarrow \quad \lambda_i = 0, \ \text{ pour tout } i \in I.$$

Une famille qui n'est pas libre est dite *liée.*

On dit que $(e_i)_{i\in I}$ est une famille *génératrice* si le sous-espace vectoriel engendré par ses éléments est E tout entier, ou autrement dit, si pour tout élément $x \in E$, il existe un sous-ensemble $J \subset I$ **fini** et une famille $(\lambda_j)_{j\in J}$ d'éléments de K tels que

$$x = \sum_{j\in J} \lambda_j e_j.$$

Autrement dit, $(e_i)_{i\in I}$ est génératrice si tout $x \in E$ s'écrit sous la forme

$$x = \sum_{i\in I} \lambda_i e_i,$$

où les λ_i sont presque tous nuls.

On dit que $(e_i)_{i\in I}$ est une *base* de E si elle est à la fois libre et génératrice. Cela revient à dire tout $x \in E$ peut s'écrire de manière **unique** sous la forme

$$x = \sum_{j\in J} \lambda_j e_j,$$

où J est un sous-ensemble fini de I, ou encore si tout $x \in E$ s'écrit de manière unique sous la forme

$$x = \sum_{i\in I} \lambda_i e_i,$$

où les $(\lambda_i)_{i\in I}$ est une famille d'éléments de K presque tous nuls.

Il est bien connu que tout espace vectoriel pouvant être engendré par un nombre fini d'éléments possède au moins une base. Le résultat suivant montre que l'on peut s'affranchir de cette hypothèse, mais sa démonstration utilise le lemme de Zorn. Comme nous travaillerons essentiellement en dimension finie, pour la plupart du temps, on pourra passer cette démonstration en première lecture.

2.2. Théorème (de la base incomplète). *Soit E un K-espace vectoriel, et soit $\mathscr{G}$ une famille génératrice de E.*

Alors, toute famille libre de E contenue dans $\mathscr{G}$ peut être complétée en une base de E formée d'éléments de $\mathscr{G}$.

En particulier, toute famille libre de E peut être complétée en une base, et tout K-espace vectoriel admet au moins une base.

Démonstration. Soit $\mathscr{L}$ une famille libre de E contenue dans $\mathscr{G}$, et soit $\mathscr{F}$ l'ensemble des familles libres $\mathscr{L}'$ de E telles que $\mathscr{L} \subset \mathscr{L}' \subset \mathscr{G}$. L'ensemble $\mathscr{F}$ est partiellement ordonné par la relation d'inclusion.

Montrons que $\mathscr{F}$ est un ensemble inductif non vide (cf. définition I-4.5). Puisque $\mathscr{L} \in \mathscr{F}$, $\mathscr{F}$ est non vide. Soit maintenant $(I, \leqslant)$ un ensemble

totalement ordonné, et $\mathscr{F}' = \{\mathscr{L}_i \mid i \in I\}$ une partie de $\mathscr{F}$ totalement ordonnée par I, c'est-à-dire

$$\text{pour tous } i, j \in I,\ i \leqslant j \quad \Longrightarrow \quad \mathscr{L}_i \subset \mathscr{L}_j.$$

La famille $\mathscr{L}' = \bigcup_{i \in I} \mathscr{L}_i$ est une famille libre de E.

En effet, soient $x_1, \ldots, x_n \in \mathscr{L}'$, et soient $\lambda_1, \ldots, \lambda_n \in K$ vérifiant

$$\lambda_1 x_1 + \cdots + \lambda_n x_n = 0.$$

Pour tout $j \in [\![1, n]\!]$, il existe $i_j \in I$ tel que $x_j \in \mathscr{L}_{i_j}$. Soit $i_0 = \max_{1 \leqslant j \leqslant n} i_j$.

Puisque la famille $(\mathscr{L}_i)_{i \in I}$ est totalement ordonnée, on a $x_1, \ldots, x_n \in \mathscr{L}_{i_0}$. La famille $\mathscr{L}_{i_0}$ étant libre, on en déduit que $\lambda_1 = \cdots = \lambda_n = 0$. Ainsi, la famille $\mathscr{L}'$ est libre, et vérifie clairement $\mathscr{L} \subset \mathscr{L}' \subset \mathscr{G}$, d'où $\mathscr{L}' \in \mathscr{F}$. Comme $\mathscr{L}_i \subset \mathscr{L}'$, pour tout $i \in I$, $\mathscr{L}'$ est alors un majorant de $\mathscr{F}'$.

Par le lemme de Zorn (théorème I-4.5), l'ensemble $\mathscr{F}$ possède un élément maximal $\mathscr{B}$. Montrons que $\mathscr{B}$ est une base de E telle que $\mathscr{L} \subset \mathscr{B} \subset \mathscr{G}$. C'est par définition une famille libre vérifiant les inclusions précédentes. Il reste donc à voir que $\mathscr{B}$ est génératrice.

Soit $x \in E$. On doit montrer que x s'écrit comme une combinaison linéaire d'éléments de $\mathscr{B}$. Commençons par le cas où $x \in \mathscr{G}$. Si $x \in \mathscr{B}$, il n'y a rien à faire. Supposons donc que $x \notin \mathscr{B}$. Si la famille $\mathscr{B} \cup \{x\}$ était libre, alors ce serait un élément de $\mathscr{F}$ contenant strictement $\mathscr{B}$, ce qui contredirait la maximalité de $\mathscr{B}$. Ainsi, la famille $\mathscr{B} \cup \{x\}$ est liée.

Il existe donc des éléments $x_1, \ldots, x_n \in \mathscr{B}$ et des scalaires $\lambda_1, \ldots, \lambda_n, \lambda \in K$ non tous nuls, tels que

$$\lambda_1 x_1 + \cdots + \lambda_n x_n + \lambda x = 0.$$

Si $\lambda = 0$, alors puisque $\mathscr{B}$ est libre, on obtient $\lambda_1 = \cdots = \lambda_n = 0$, d'où une contradiction. Ainsi, $\lambda \neq 0$. On a alors

$$x = -\lambda^{-1}(\lambda_1 x_1 + \cdots + \lambda_n x_n),$$

et x s'écrit comme une combinaison linéaire d'éléments de $\mathscr{B}$. On a donc démontré que tout élément de $\mathscr{G}$ s'écrit comme une combinaison linéaire d'éléments de $\mathscr{B}$. Mais, $\mathscr{G}$ étant génératrice, tout élément de E s'écrit comme combinaison linéaire d'éléments de $\mathscr{G}$. Il s'ensuit aisément que tout élément de E est alors combinaison linéaire d'éléments de $\mathscr{B}$. Ainsi, $\mathscr{B}$ est génératrice, et est donc bien une base.

La dernière partie du théorème s'obtient en appliquant ce qui précède à la famille libre vide et $\mathscr{G} = E$. Cela achève la démonstration. □

2.3. Remarque. L'espace vectoriel nul admet une base, à savoir la base vide.

Rappelons le résultat suivant.

2.4. Théorème. *Soit $u \in \mathscr{L}(E,F)$. Alors, les propriétés suivantes sont équivalentes :*

(1) *l'application u est injective, (resp. surjective, resp. bijective) ;*

(2) *pour toute base $(e_i)_{i\in I}$, la famille $\big(u(e_i)\big)_{i\in I}$ est libre (resp. génératrice, resp. une base) ;*

(3) *il existe une base $(e_i)_{i\in I}$ telle que la famille $\big(u(e_i)\big)_{i\in I}$ est libre (resp. génératrice, resp. une base).*

On a aussi le résultat suivant.

2.5. Proposition. *Soient E et F deux K-espaces vectoriels, et soit $(e_i)_{i\in I}$ une base de E. Alors, pour toute famille $(y_i)_{i\in I}$ d'éléments de F, il existe une unique application linéaire $u : E \longrightarrow F$ telle que*

$$u(e_i) = y_i, \qquad \text{pour tout } i \in I.$$

On rappelle maintenant la théorie de la dimension. Si X est un ensemble, on note $|X|$ son nombre d'éléments (éventuellement infini).

2.6. Théorème. *Soit E un K-espace vectoriel. Supposons que E admette une famille génératrice finie. Alors, toutes les bases de E sont finies, et ont le même nombre d'éléments, disons n. De plus, pour toute famille libre $\mathscr{L}$ et toute famille génératrice $\mathscr{G}$, on a*

$$|\mathscr{L}| \leqslant n \leqslant |\mathscr{G}|.$$

2.7. Définition. Un K-espace vectoriel E est dit de *type fini* s'il admet une famille génératrice finie. Dans ce cas, la *dimension* de E est le cardinal de n'importe quelle base de E. Elle est notée $\dim_K(E)$.

2.8. Théorème. *Soient $E, E_1, \ldots, E_n, F$ des K-espaces vectoriels de dimension finie. Alors :*

(1) *on a $\dim_K(K^n) = n$, pour tout $n \geqslant 1$;*

(2) *si $\dim_K(E) = n$, alors $E \simeq K^n$;*

(3) *deux K-espaces vectoriels de dimension finie sont isomorphes si, et seulement si, ils ont même dimension ;*

(4) *pour tout sous-espace vectoriel E' de E, E' est de dimension finie, et l'on a*
$$\dim_K(E') \leqslant \dim_K(E).$$
De plus, on a égalité si, et seulement si, $E' = E$;

(5) *le K-espace vectoriel $E_1 \times \cdots \times E_n$ est de dimension finie, et l'on a*

$$\dim_K(E_1 \times \cdots \times E_n) = \sum_{i=1}^{n} \dim_K(E_i) \ ;$$

(6) *le K-espace vectoriel $\mathscr{L}(E,F)$ est de dimension finie, et l'on a*

$$\dim_K \big(\mathscr{L}(E,F)\big) = \dim_K(E) \times \dim_K(F) \ ;$$

(7) Théorème du rang *: pour tout $u \in \mathscr{L}(E,F)$, on a*

$$\dim_K(E) = \dim_K \big(\mathrm{Ker}(u)\big) + \dim_K \big(\mathrm{Im}(u)\big).$$

2.9. Corollaire. *Soient E, F deux K-espaces vectoriels de dimension finie. On suppose que $\dim_K(E) = \dim_K(F)$. Alors, pour tout $u \in \mathscr{L}(E,F)$ les propriétés suivantes sont équivalentes :*

(1) *l'application u est injective ;*

(2) *l'application u est surjective ;*

(3) *l'application u est bijective.*

2.10. Corollaire. *Soit E un K-espace vectoriel de dimension n, et soit* **e** *une famille de n éléments de E. Alors, les propriétés suivantes sont équivalentes :*

(1) *la famille* **e** *est libre ;*

(2) *la famille* **e** *est génératrice ;*

(3) *la famille* **e** *est une base de E.*

3. Somme d'espaces vectoriels

3.1. Définition. Soit E un K-espace vectoriel, et soit $(E_i)_{i\in I}$ une famille de sous-espaces vectoriels de E. On dit que les sous-espaces E_i, $i \in I$ sont en *somme directe* si pour tout sous-ensemble fini $J \subset I$ et tout $x_j \in E_j$, $j \in J$, on a

$$\sum_{j\in J} x_j = 0 \Longrightarrow x_j = 0, \quad \text{pour tout } j \in J.$$

La *somme* $\sum_{i\in I} E_i$ des sous-espaces E_i, $i \in I$ est le sous-espace vectoriel engendré par la réunion des E_i. On vérifie facilement que l'on a

$$\sum_{i\in I} E_i = \Big\{ \sum_{j\in J} x_j \mid J \subset I \text{ fini }, \ x_j \in J \Big\}.$$

En particulier, on a

$$E_1 + \cdots + E_n = \big\{ x_1 + \cdots + x_n \mid x_i \in E_i, \ i \in [\![1,n]\!] \big\}.$$

On dit que E est la *somme directe* des E_i, $i \in I$, si :

(1) on a $E = \sum_{i \in I} E_i$;

(2) les sous-espaces E_i, $i \in I$ sont en somme directe.

On le note $E = \bigoplus_{i \in I} E_i$. Autrement dit, $E = \bigoplus_{i \in I} E_i$ si, et seulement si, tout élément de E s'écrit de manière *unique* comme une somme finie d'éléments des E_i, $i \in I$.

Dans ce cas, on obtient une base E en recollant des bases de chaque E_i.

Par exemple, si $(e_i)_{i \in I}$ est une base de E, on a $E = \bigoplus_{i \in I} Ke_i$.

On montre facilement que $E = E_1 \oplus E_2$ si, et seulement si, $E = E_1 + E_2$ et $E_1 \cap E_2 = \{0\}$.

La proposition suivante généralise la proposition 2.5.

3.2. Proposition. *Soit E un K-espace vectoriel, et soit $(E_i)_{i \in I}$ une famille de sous-espaces vectoriels de E tels que $E = \bigoplus_{i \in I} E_i$.*

Pour tout K-espace vectoriel F, et tout $f_i \in \mathscr{L}(E_i, F)$, il existe une unique application linéaire $f \in \mathscr{L}(E, F)$ telle que $f_{|E_i} = f_i$, pour tout $i \in I$.

Enfin, la dimension d'une somme directe se calcule facilement.

3.3. Proposition. *Soit K un corps, et soit E un K-espace vectoriel de dimension finie.*

Si $E = E_1 \oplus \cdots \oplus E_n$, alors $\dim_K(E) = \dim_K(E_1) + \cdots + \dim_K(E_n)$.

3.4. Définition. Soit E un K-espace vectoriel, et soit F un sous-espace vectoriel de E. On dit qu'un sous-espace vectoriel G de E est un *supplémentaire* de F dans E si l'on a $E = F \oplus G$.

3.5. Proposition. *Soit E un K-espace vectoriel. Alors, tout sous-espace vectoriel de E admet un supplémentaire.*

Démonstration. Soit F un sous-espace vectoriel de E, et soit $(f_i)_{i \in I}$ une base de F. En particulier, c'est une famille libre de E. Elle peut donc être complétée une base de E. Il existe donc une famille $(f_j)_{j \in J}$ d'éléments de E telle que $(f_k)_{k \in I \cup J}$ soit une base de E (avec I et J disjoints). On montre alors facilement que $G = \mathrm{Vect}_K(f_j,\ j \in J)$ est le supplémentaire recherché. □

4. Matrices représentatives, changement de base

Soit E un K-espace vectoriel de dimension n, et soit $\mathbf{e} = (e_1, \ldots, e_n)$ une base de E. Si $x = x_1 e_1 + \cdots + x_n e_n$, on note

$$[x]_{\mathbf{e}} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$$

le vecteur des coordonnées de x dans la base $\mathbf{e}$.

Alors, l'application

$$\begin{aligned} \varphi_{\mathbf{e}} : \quad K^n &\longrightarrow E \\ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} &\longmapsto \sum_{i=1}^{n} x_i e_i \end{aligned}$$

est un isomorphisme de K-espaces vectoriels, d'inverse

$$\begin{aligned} \psi_{\mathbf{e}} : E &\longrightarrow K^n \\ x &\longmapsto [x]_{\mathbf{e}}. \end{aligned}$$

4.1. Définition. Soient E, F deux K-espaces vectoriels non nuls de dimensions n et m respectivement, et soit $u \in \mathscr{L}(E, F)$. Soit $\mathbf{e}$ une base de E et soit $\mathbf{f}$ une base de F. La *matrice représentative* de u dans les bases $\mathbf{e}$ et $\mathbf{f}$ est la matrice dont les colonnes sont les vecteurs

$$[u(e_1)]_{\mathbf{f}}, \ \ldots, \ [u(e_n)]_{\mathbf{f}}.$$

On la note $\mathrm{Mat}(u\,; \mathbf{e}, \mathbf{f})$. Si $u(e_j) = \sum_{i=1}^{m} a_{ij} f_i$ pour $j \in [\![1, n]\!]$, alors on a

$$\mathrm{Mat}(u\,; \mathbf{e}, \mathbf{f}) = (a_{ij}) \in \mathrm{M}_{m \times n}(K).$$

Si $E = F$ et $\mathbf{e} = \mathbf{f}$, on l'appelle *matrice représentative* de u dans la base $\mathbf{e}$ et on la note $\mathrm{Mat}(u\,; \mathbf{e})$.

4.2. Lemme. *Soient E, F deux K-espaces vectoriels de dimension finie n et m respectivement. Soit $u \in \mathscr{L}(E, F)$. Soit $\mathbf{e}$ une base de E et soit $\mathbf{f}$ une base de F. Alors, pour tout $x \in E$, on a*

$$[u(x)]_{\mathbf{f}} = \mathrm{Mat}(u\,; \mathbf{e}, \mathbf{f})[x]_{\mathbf{e}}.$$

4.3. Proposition. *Soient E_1, E_2, E_3 trois K-espaces vectoriels de dimension finie, soient $u \in \mathscr{L}(E_1, E_2)$, $v \in \mathscr{L}(E_2, E_3)$, et soient $\mathbf{e}_1$, $\mathbf{e}_2$, $\mathbf{e}_3$ des bases de E_1, E_2, E_3 respectivement. Alors,*

$$\mathrm{Mat}(v \circ u\,; \mathbf{e}_1, \mathbf{e}_3) = \mathrm{Mat}(v\,; \mathbf{e}_2, \mathbf{e}_3)\mathrm{Mat}(u\,; \mathbf{e}_1, \mathbf{e}_2).$$

En particulier, $u \in \mathscr{L}(E, F)$ est inversible si, et seulement si, sa matrice représentative dans des bases arbitraires de E et F est inversible.

4.4. Définition. Soient $\mathbf{e}$, $\mathbf{e}'$ deux bases d'un K-espace vectoriel E de dimension $n \geqslant 1$. La *matrice de passage* de $\mathbf{e}'$ à $\mathbf{e}$ est la matrice dont les colonnes sont les vecteurs

$$[e'_1]_{\mathbf{e}}, \ \ldots, \ [e'_n]_{\mathbf{e}},$$

c'est-à-dire la matrice $\mathrm{Mat}(\mathrm{Id}_E\,; \mathbf{e}', \mathbf{e})$. En particulier, c'est une matrice inversible de $\mathrm{M}_n(K)$. Son inverse est la matrice de passage de $\mathbf{e}$ à $\mathbf{e}'$.

Attention ! La plupart des auteurs définissent la matrice de passage de la base $\mathbf{e}'$ à la base $\mathbf{e}$ comme la matrice dont les colonnes sont les vecteurs

$$[e_1]_{\mathbf{e}'}, \ \ldots, \ [e_n]_{\mathbf{e}'}.$$

Néanmoins, nous préférons notre définition, plus cohérente avec le résultat ci-dessous.

4.5. Lemme. *Soient $\mathbf{e}$, $\mathbf{e}'$ deux bases de E, et soit P la matrice de passage de $\mathbf{e}'$ à $\mathbf{e}$. Alors, pour tout $x \in E$, on a*

$$[x]_{\mathbf{e}} = P[x]_{\mathbf{e}'}.$$

4.6. Remarque. Appliquer la matrice de passage de $\mathbf{e}'$ à $\mathbf{e}$ permet donc de passer des coordonnées dans la base $\mathbf{e}'$ aux coordonnées dans la base $\mathbf{e}$.

4.7. Proposition. *Soit E un espace vectoriel de dimension finie sur un corps K, et soit $u \in \mathscr{L}(E)$. Soient $\mathbf{e}$ et $\mathbf{e}'$ deux bases de E, et soit P la matrice de passage de $\mathbf{e}'$ à $\mathbf{e}$. Alors, on a*

$$P\,\mathrm{Mat}(u\,; \mathbf{e}')P^{-1} = \mathrm{Mat}(u\,; \mathbf{e}).$$

Nous finissons ce paragraphe en introduisant la notion de sous-espace stable, et qui sera cruciale dans le problème de la réduction des endomorphismes.

4.8. Définition. Soit $u \in \mathscr{L}(E)$. Un sous-espace vectoriel F de E est dit *stable* par u si $u(F) \subset F$. Dans ce cas, u induit un endomorphisme de F

$$\begin{aligned} u_F : F &\longrightarrow F \\ x &\longmapsto u(x). \end{aligned}$$

4.9. Remarque. L'intérêt d'avoir un sous-espace stable F est que l'on peut l'utiliser pour trouver une base de E dans laquelle la matrice représentative est triangulaire par blocs.

En effet, soit $\mathbf{f} = (f_1, \ldots, f_d)$ une base de F, et complétons-la en une base de E, disons $\mathbf{e}$. Puisque F est stable par u, on a $u(f_i) \in F$, $i \in [\![1, d]\!]$, et donc $u(f_i)$ est combinaison linéaire de $f_1, \ldots, f_d$. On a ainsi

$$\mathrm{Mat}(u\,;\mathbf{e}) = \begin{pmatrix} \mathrm{Mat}(u_F\,;\mathbf{f}) & B \\ 0 & C \end{pmatrix}.$$

Si maintenant on a $E = E_1 \oplus \cdots \oplus E_r$, où chaque E_i est stable par u, et si l'on note $\mathbf{e}$ la base de E obtenue en recollant des bases $\mathbf{e}_i$ de E_i, alors on a

$$\mathrm{Mat}(u\,;\mathbf{e}) = \begin{pmatrix} \mathrm{Mat}(u_{E_1}\,;\mathbf{e}_1) & & \\ & \ddots & \\ & & \mathrm{Mat}(u_{E_r}\,;\mathbf{e}_r) \end{pmatrix}.$$

Cela sera appliqué de manière systématique dans la cinquième partie, lorsque nous étudierons la réduction des endomorphismes en détail.

5. Formes linéaires, droites et hyperplans

Nous finissons ce chapitre en faisant de brefs rappels sur la dualité. Celle-ci sera étudiée de manière approfondie dans le chapitre XXXII.

5.1. Définition. Soit E un K-espace vectoriel.

Une *forme linéaire* sur E est un élément de $\mathscr{L}(E, K)$, autrement dit une application linéaire $\varphi : E \longrightarrow K$. Le *dual* de E, noté E^*, est le K-espace vectoriel $\mathscr{L}(E, K)$.

5.2. Exemples

(1) Pour tout $n \geqslant 0$, l'application

$$\begin{array}{ccc} \mathbb{R}[X] & \longrightarrow & \mathbb{R} \\ P & \longmapsto & \dfrac{P^{(n)}(0)}{n!} \end{array}$$

est une forme linéaire.

(2) L'application

$$\begin{array}{ccc} C^0([0,1], \mathbb{R}) & \longrightarrow & \mathbb{R} \\ f & \longmapsto & \displaystyle\int_0^1 f(t)\mathrm{d}t \end{array}$$

est une forme linéaire.

(3) Soit E un K-espace vectoriel, et soit $(e_i)_{i\in I}$ une base de E. D'après la proposition 2.5, pour tout indice $i \in I$, il existe une unique forme linéaire $e_i^* : E \longrightarrow K$ telle que

$$e_i^*(e_k) = \delta_{ik}, \quad \text{pour tout } k \in I.$$

Elle est simplement définie par

$$\begin{aligned} e_i^* : \quad E \quad &\longrightarrow K \\ \sum_{k\in I} x_k e_k &\longmapsto x_i. \end{aligned}$$

Par définition, on a alors

$$x = \sum_{i\in I} e_i^*(x) e_i, \quad \text{pour tout } x \in E.$$

5.3. Définition. Soit E un K-espace vectoriel, et soit $\mathbf{e} = (e_i)_{i\in I}$ une base de E. Les formes linéaires e_i^*, $i \in I$ sont appelées les *formes coordonnées* par rapport à la base $\mathbf{e}$.

Ces formes coordonnées jouent un rôle important pour l'étude des formes linéaires en dimension finie, au vu de la proposition suivante.

5.4. Proposition. *Soit E un K-espace vectoriel, et soit $\mathbf{e} = (e_i)_{i\in I}$ une base de E. La famille $\mathbf{e}^* = (e_i^*)_{i\in I}$ est une famille libre de E^*.*

De plus, $\mathbf{e}^*$ est une base de E^* si, et seulement si, E est de dimension finie. Dans ce cas, pour tout $\varphi \in E^*$, on a

$$x = \sum_{i\in I} e_i^*(x) e_i \quad \text{et} \quad \varphi = \sum_{i\in I} \varphi(e_i) e_i^*.$$

Enfin, E^* est de dimension finie si, et seulement si, E est de dimension finie. Dans ce cas, on a

$$\dim_K(E^*) = \dim_K(E).$$

Démonstration. Montrons que $\mathbf{e}^*$ est libre. Soit $(\lambda_i)_{i\in I}$ une famille d'éléments de K presque tous nuls, et supposons que

$$\sum_{i\in I} \lambda_i e_i^* = 0_{E^*}.$$

Soit $k \in I$. En appliquant e_k à cette égalité, par définition des formes coordonnées, on obtient $\lambda_k = 0$, d'où le résultat.

Si E est de dimension infinie, alors I est infini, et $\mathbf{e}^*$ est une famille libre de E de cardinal infini. Ainsi, E est de dimension infinie. Si E est de

dimension finie n, alors $E^* = \mathscr{L}(E,K)$ est de dimension finie. De plus,

$$\dim_K(E^*) = \dim_K\big(\mathscr{L}(E,K)\big) = \dim_K(E)\times\dim_K(K) = \dim_K(E) = n.$$

Dans ce cas, comme $\mathbf{e}^*$ est libre de cardinal n, c'est une base de E^*.

Supposons que E soit de dimension finie, et soit $\varphi \in E^*$. Pour tout $k \in I$, on a

$$\Big(\sum_{i\in I}\varphi(e_i)e_i^*\Big)(e_k) = \sum_{i\in I}\varphi(e_i)e_i^*(e_k) = \varphi(e_k).$$

Ainsi, φ et $\sum\limits_{i\in I}\varphi(e_i)e_i^*$ coïncident sur une base, et elle sont donc égales. Remarquons que cette égalité n'a de sens pour tout $\varphi \in E^*$ que si I est fini, et donc que si E est de dimension finie. En effet, si E est de dimension infinie, il n'y a aucune raison que pour les scalaires $\varphi(e_i)$, $i \in I$ soient presque tous nuls. C'est d'ailleurs sur cette observation que nous allons démontrer que $\mathbf{e}^*$ n'est pas une base de E^* si E est de dimension infinie.

Soit $W = \mathrm{Vect}_K(e_i^*,\, i \in I)$. Remarquons que si $\varphi \in W$, alors il existe un sous-ensemble fini J de I tel que

$$\varphi = \sum_{j\in J}\lambda_j e_j^*.$$

En particulier, pour tout $k \in I \setminus J$, on a $\varphi(e_k) = 0$. Il s'ensuit que si E est de dimension infinie, l'unique forme linéaire $\varphi \in E^*$ vérifiant

$$\varphi(e_i) = 1, \quad \text{pour tout } i \in I$$

n'est pas un élément de W. Ainsi, si E est de dimension infinie, $W \neq E^*$, et donc $\mathbf{e}^*$ n'est pas génératrice. Cela achève la démonstration. □

5.5. Définition. Soit E un K-espace vectoriel de dimension finie $n \geqslant 1$, et soit $\mathbf{e} = (e_1,\dots,e_n)$ une base de E. La base $\mathbf{e}^* = (e_1^*,\dots,e_n^*)$ de E^* est appelée la *base duale* de $\mathbf{e}$.

Les formes linéaires jouent un rôle très important en géométrie, car elles sont intimement liées aux hyperplans, comme nous allons l'expliquer maintenant. On commence par rappeler une définition.

5.6. Définition. Soit E un K-espace vectoriel non nul. Une *droite* de E est un sous-espace vectoriel D de dimension 1, c'est-à-dire un sous-espace D de la forme

$$D = Kx_0, \quad x_0 \in E \setminus \{0\}.$$

Un *hyperplan* de E est un sous-espace vectoriel H qui est le supplémentaire d'une droite D. Autrement dit, un sous-espace H est un hyperplan s'il existe une droite D de E telle que $E = H \oplus D$.

5.7. Remarque. Une droite ou un hyperplan de E n'existent que si E est non nul.

Le lemme suivant caractérise les hyperplans.

5.8. Lemme. *Soit E est un K-espace vectoriel non nul, et soit H un sous-espace de E. Alors, les propriétés suivantes sont équivalentes :*

(1) *le sous-espace H est un hyperplan ;*

(2) *le sous-espace H est le noyau d'une forme linéaire $\varphi : E \longrightarrow K$ non nulle.*

Dans ce cas, on a $E = H \oplus Kx_0$, où $x_0 \in E$ vérifie $\varphi(x_0) = 1$.

De plus, une forme linéaire $\psi : E \longrightarrow K$ s'annule sur H si, et seulement si, il existe $\lambda \in K$ tel que $\psi = \lambda\varphi$.

(3) *Enfin, si E est de dimension finie, ces deux conditions sont aussi équivalentes à :*

$\dim_K(H) = \dim_K(E) - 1$.

Démonstration. Soit H un hyperplan de E, et soit $D = Kx_0$ une droite de E telle que $E = H \oplus D$. Alors, l'application

$$\begin{array}{rccc} \varphi : & E & \longrightarrow & K \\ & x_H + \lambda x_0 & \longmapsto & \lambda \end{array}$$

est une forme linéaire non nulle (puisque $\varphi(x_0) = 1$) de noyau H. Remarquons aussi que si E est de dimension finie, on a

$$\dim_K(E) = \dim_K(H) + \dim_K(D) = \dim_K(H) + 1,$$

et ainsi $\dim_K(H) = \dim_K(E) - 1$.

Soit maintenant $\psi : E \longrightarrow K$ une forme linéaire. Si $\psi = \lambda\varphi$, alors ψ s'annule sur $\mathrm{Ker}(\varphi) = H$. Inversement, supposons que ψ s'annule sur H. Pour tout $x_H \in H$, et tout $\lambda \in K$, on a

$$\psi(x_H + \lambda x_0) = \lambda\psi(x_0) = \psi(x_0)\varphi(x_H + \lambda x_0).$$

Ainsi, $\psi = \psi(x_0)\varphi$.

Cela démontre (1) $\Longrightarrow$ (2), et aussi (1) $\Longrightarrow$ (3) lorsque E est de dimension finie. Supposons maintenant que H soit le noyau d'une forme linéaire non nulle $\varphi : E \longrightarrow K$. Remarquons tout d'abord qu'il existe un élément $x_0 \in E$ tel que $\varphi(x_0) = 1$. En effet, soit $y_0 \in E$ un vecteur tel que $\varphi(y_0) \neq 0$, et posons $x_0 = \dfrac{y_0}{\varphi(y_0)}$. Alors, on a

$$\varphi(x_0) = \frac{\varphi(y_0)}{\varphi(y_0)} = 1.$$

Fixons maintenant un tel $x_0 \in E$. Remarquons que $x_0 \neq 0$ (car sinon, on aurait $\varphi(x_0) = \varphi(0) = 0$).

Nous allons montrer que $E = H \oplus Kx_0$, ce qui prouvera que H est un hyperplan. Soit $x \in E$. Alors, $x_H = x - \varphi(x)x_0 \in H$. En effet, on a

$$\varphi(x_H) = \varphi\big(x - \varphi(x)x_0\big) = \varphi(x) - \varphi(x)\varphi(x_0) = \varphi(x) - \varphi(x) = 0.$$

Ainsi, $x_H \in \mathrm{Ker}(\varphi) = H$. On a alors $x = x_H + \varphi(x)x_0$, d'où $E = H + Kx_0$. Montrons que la somme est directe. Si $x \in H \cap Kx_0$, alors il existe $\lambda \in K$ tel que $x = \lambda x_0$. Comme $x \in H = \mathrm{Ker}(\varphi)$, on a donc

$$0 = \varphi(x) = \varphi(\lambda x_0) = \lambda\varphi(x_0) = \lambda.$$

Il s'ensuit que $x = 0$, d'où le résultat.

Il reste à vérifier que (3) $\Longrightarrow$ (1). Soit $(x_1, \ldots, x_{n-1})$ une base de H. Choisissons $x_0 \in E$ tel que $(x_0, \ldots, x_{n-1})$ soit une base de E. Alors, $E = H \oplus Kx_0$, et H est bien un hyperplan de E. Cela achève la démonstration. □

5.9. Remarque. Ce résultat se retraduit de manière peut-être plus intuitive lorsque $E = K^n$. Dans ce cas, ce lemme ne dit rien d'autre que la chose suivante : un sous-espace H de K^n est un hyperplan si, et seulement si, H est l'ensemble des solutions d'une équation

$$a_1x_1 + \cdots + a_nx_n = 0, \quad x_i \in K,$$

où les a_i sont non tous nuls.

De plus, deux équations définissant un même hyperplan H sont nécessairement proportionnelles.

Le lemme qui suit est une généralisation de la remarque précédente.

5.10. Lemme. *Soit K un corps, et soit E un K-espace vectoriel de dimension $n \geqslant 0$. Soit $d \in [\![0, n]\!]$. Pour toutes formes linéaires $\varphi_1, \ldots, \varphi_d \in E^*$ linéairement indépendantes, on a*

$$\dim_K \Big(\bigcap_{i=1}^{d} \mathrm{Ker}(\varphi_i) \Big) = n - d.$$

Démonstration. Soit $\mathbf{e} = (e_1, \ldots, e_n)$ une base de E. Remarquons tout d'abord que les lignes $L_1, \ldots, L_d \in K^n$ de la matrice

$$M = \big(\varphi_i(e_j)\big) \in \mathrm{M}_{d\times n}(K)$$

sont linéairement indépendantes.

En effet, soient $\lambda_1, \dots, \lambda_d \in K$ tels que $\lambda_1 L_1 + \cdots + \lambda_d L_d = 0$. Alors, pour tout $j \in [\![1, n]\!]$, on a

$$(\lambda_1\varphi_1 + \cdots + \lambda_d\varphi_d)(e_j) = \lambda_1\varphi_1(e_j) + \cdots + \lambda_d\varphi_d(e_j) = 0.$$

Ainsi, $\lambda_1\varphi_1 + \cdots + \lambda_d\varphi_d$ est nulle sur une base, donc identiquement nulle. Par hypothèse, on en déduit que $\lambda_1 = \cdots = \lambda_d = 0$.

Par suite, M est de rang d. Or, pour tout $x \in E$, $M[x]_{\mathbf{e}} = \begin{pmatrix} \varphi_1(x) \\ \vdots \\ \varphi_d(x) \end{pmatrix}$. On en conclut aisément que $\mathrm{Ker}(M)$ est isomorphe à $\bigcap_{i=1}^{d} \mathrm{Ker}(\varphi_i)$. Mais alors,

$$\dim_K\Big(\bigcap_{i=1}^{d} \mathrm{Ker}(\varphi_i)\Big) = \dim_K\big(\mathrm{Ker}(M)\big) = n - \mathrm{rg}(M) = n - d,$$

d'où le résultat voulu. □

5.11. Remarque. Si $E = K^n$, ce résultat se retraduit comme suit : soit un système linéaire homogène de d équations à n inconnues. Si la matrice du système linéaire est de rang d, l'ensemble des solutions du système linéaire est de dimension $n - d$.

Nous approfondirons ce point de vue et l'étude de la dualité en dimension quelconque dans le chapitre XXXII. En particulier, des exercices sur les formes linéaires ne seront pas proposés dans ce chapitre.

6. Exercices

E1. Soit E un K-espace vectoriel, et soient F, G deux sous-espaces de dimension finie. En appliquant le théorème du rang à l'application

$$\begin{aligned} F \times G &\longrightarrow E \\ (x_F, x_G) &\longmapsto x_F + x_G, \end{aligned}$$

montrer la formule de Grassmann :

$$\dim_K(F + G) = \dim_K(F) + \dim_K(G) - \dim_K(F \cap G).$$

E2. Soit E un K-espace vectoriel de dimension finie $n \geqslant 0$. Soient F et G deux sous-espaces vectoriels de E tels que $\dim_K(F) = \dim_K(G) = k$.

Montrer par récurrence sur $n - k$ que F et G ont un supplémentaire commun.

Indication. Si $F = G$, il n'y a rien à faire. Si $F \neq G$, justifier que les ensembles $F \setminus G$ et $G \setminus F$ sont non vides. Choisir $x \in F \setminus G$ et $x' \in G \setminus F$. Montrer alors que $x + x' \notin F \cup G$ et considérer $F' = F + \mathrm{Vect}_K(x + x')$ et $G' = G + \mathrm{Vect}_K(x + x')$.

E3. Soit $(e_1, \ldots, e_p)$ une famille libre de vecteurs d'un K-espace vectoriel E, soit $F = \mathrm{Vect}_K(e_1, \ldots, e_p)$ et soit G un supplémentaire de F dans E. Pour tout $a \in G$, on note

$$F_a = \mathrm{Vect}_K(e_1 + a, \ldots, e_p + a)$$

a. Montrer que pour tout $a \in G$, on a $E = F_a \oplus G$.

b. Montrer que pour tous $a, b \in G$, on a $a \neq b \implies F_a \neq F_b$.

c. En déduire que si K est infini, G possède une infinité de supplémentaires.

E4. On se propose ici de déterminer tous les automorphismes de K-algèbre de $\mathrm{M}_n(K)$, c'est-à-dire les applications K-linéaires

$$\theta : \mathrm{M}_n(K) \longrightarrow \mathrm{M}_n(K)$$

bijectives telles que $\theta(AB) = \theta(A)\theta(B)$ pour tous $A, B \in \mathrm{M}_n(K)$ et $\theta(\mathrm{I}_n) = \mathrm{I}_n$.

a. Soit $P \in \mathrm{GL}_n(K)$ une matrice inversible. Montrer que l'application

$$\begin{aligned} \mathrm{Int}(P) : \mathrm{M}_n(K) &\longrightarrow \mathrm{M}_n(K) \\ A &\longmapsto PAP^{-1} \end{aligned}$$

est un automorphisme de K-algèbre de $\mathrm{M}_n(K)$.

Le but de la suite est de démontrer la réciproque.

On note $E_{ij} \in \mathrm{M}_n(K)$ la matrice dont tous les coefficients sont nuls, sauf celui en position (i, j), qui vaut 1.

Soit $\theta : \mathrm{M}_n(K) \longrightarrow \mathrm{M}_n(K)$ un automorphisme de K-algèbre. On pose alors $F_{ij} = \theta(E_{ij})$

b. Vérifier que, pour tous $i, j, k, \ell \in [\![1, n]\!]$, on a

$$F_{ij}F_{k\ell} = 0 \text{ si } j \neq k, \quad F_{ij}F_{j\ell} = F_{i\ell},$$

ainsi que

$$F_{11} + \cdots + F_{nn} = \mathrm{I}_n.$$

c. Montrer que $K^n = \mathrm{Im}(F_{11}) \oplus \cdots \oplus \mathrm{Im}(F_{nn})$.

d. Montrer que l'application

$$\begin{aligned} K^n &\longrightarrow K^n \\ X &\longmapsto F_{i,1}X \end{aligned}$$

induit par double restriction un isomorphisme de $\mathrm{Im}(F_{11})$ sur $\mathrm{Im}(F_{ii})$.

e. Déduire des questions précédentes que l'on a $\dim_K\left(\mathrm{Im}(F_{ii})\right) = 1$ pour tout $i \in [\![1, n]\!]$, et que si X_1 est un vecteur non nul de $\mathrm{Im}(F_{11})$, alors la famille de vecteurs $(F_{11}X_1, \ldots, F_{n1}X_1)$ est une base de K^n.

f. Soit $P \in \mathrm{M}_n(K)$ la matrice dont la j-ème colonne est $F_{j1}X_1$. Justifier que $P \in \mathrm{GL}_n(K)$.

g. Si ε_k désigne le k-ième vecteur colonne de la base canonique de K^n, montrer que $F_{ij}P\varepsilon_k = PE_{ij}\varepsilon_k$ pour tous $i, j, k \in \in [\![1, n]\!]$.

h. En déduire que $F_{ij}P = PE_{ij}$ pour tous $i, j \in [\![1, n]\!]$, puis $\theta = \mathrm{Int}(P)$.

E5. Soit E un K-espace vectoriel (non nécessairement de dimension finie), et soit $u \in \mathscr{L}(E)$. On suppose que pour tout $x \in E$, la famille $(x, u(x))$ est liée.

Montrer que u est une homothétie, i.e. il existe $\lambda \in K$ tel que $u = \lambda \mathrm{Id}_E$.

E6. Si $A = (a_{ij})_{i,j} \in \mathrm{M}_n(K)$, la *trace* de A, notée $\mathrm{tr}(A)$, est la somme des coefficients diagonaux de A, i.e.

$$\mathrm{tr}(A) = \sum_{i=1}^{n} a_{ii}.$$

a. Montrer que l'application $\mathrm{tr} : \mathrm{M}_n(K) \longrightarrow K$ est K-linéaire.

b. Montrer que pour tous $A, B \in \mathrm{M}_n(K)$, on a

$$\mathrm{tr}(A^t) = \mathrm{tr}(A) \quad \text{et} \quad \mathrm{tr}(AB) = \mathrm{tr}(BA).$$

c. Montrer que deux matrices semblables ont même trace.

d. Si E est un K-espace vectoriel de dimension finie, et si $u \in \mathscr{L}(E)$, la trace d'une matrice représentative de u est indépendante de la base choisie. On note alors $\mathrm{tr}(u)$ la trace d'une matrice représentative de u dans une base arbitraire.

e. Montrer que l'application $\mathrm{tr} : \mathscr{L}(E) \longrightarrow K$ est K-linéaire, que l'on a

$$\mathrm{tr}(u \circ v) = \mathrm{tr}(v \circ u) \quad \text{pour tous } u, v \in \mathscr{L}(E),$$

et que deux endomorphismes semblables ont même trace.

Les deux exercices suivants utilisent les résultats de l'exercice 4.6.

E7. Soit K un sous-corps de $\mathbb{C}$. Le but de cet exercice est de démontrer que tout matrice $A \in \mathrm{M}_n(K)$ de trace nulle est semblable à une matrice dont la diagonale est nulle [1]. On procède par récurrence sur n.

a. Traiter le cas $n = 1$.

On suppose le résultat vrai pour les matrices de taille n.

Soit $A \in \mathrm{M}_{n+1}(K)$ telle que $\mathrm{tr}(A) = 0$. On suppose $A \neq 0$, sinon le résultat est clair.

1. Le lecteur connaissant la notion de caractéristique d'un anneau se convaincra que les arguments présentés fonctionnent encore dans le cas d'un corps de caractéristique nulle.

b. On note u l'endomorphisme de K^{n+1} dont la matrice dans la base canonique est A. En utilisant l'exercice 5, montrer qu'il existe $x \in K^{n+1}$ tel que $(x, u(x))$ est libre.

c. En déduire qu'il existe une base $\mathbf{e}$ de K^{n+1} telle que

$$\mathrm{Mat}(u; \mathbf{e}) = \left(\begin{array}{c|c} 0 & L \\ \hline 1 & \\ 0 & \\ \vdots & M \\ 0 & \end{array}\right),$$

où $L \in \mathrm{M}_{1\times n}(L)$ et $M \in \mathrm{M}_n(K)$.

d. Justifier qu'il existe $Q \in \mathrm{GL}_n(K)$ tel $M' = Q^{-1}MQ$ soit à diagonale nulle.

e. En déduire successivement que $\mathrm{Mat}(u; \mathbf{e})$, puis A, est semblable à une matrice à diagonale nulle et conclure.

f. Montrer que le résultat est faux si K est arbitraire (prendre $K = \mathbb{F}_2$).

E8. Un *commutateur* de $\mathrm{M}_n(K)$ est une matrice de la forme

$$[A, B] = AB - BA,$$

où $A, B \in \mathrm{M}_n(K)$.

a. Montrer que tout commutateur de $\mathrm{M}_n(K)$ est de trace nulle.

b.

(i) Soit $D \in \mathrm{M}_n(K)$ une matrice diagonale dont les éléments diagonaux sont deux à deux distincts. Montrer que les matrices commutant avec D sont exactement les matrices diagonales.

Ce résultat est-il vrai pour une matrice diagonale quelconque ?

(ii) Soit $\mathscr{N}$ l'ensemble des matrices à diagonale nulle. Montrer que $\mathscr{N}$ est un sous-espace vectoriel de $\mathrm{M}_n(K)$.

(iii) Soit $D \in \mathrm{M}_n(K)$ une matrice diagonale. Montrer que pour toute matrice $N \in \mathscr{N}$, on a $[D, N] \in \mathscr{N}$, et que l'application

$$\begin{aligned} \varphi\colon \mathscr{N} &\longrightarrow \mathscr{N} \\ N &\longmapsto [D, N] \end{aligned}$$

est un endomorphisme de $\mathscr{N}$.

Montrer de plus que, si les éléments diagonaux de D sont deux à deux distincts, φ est un automorphisme de $\mathscr{N}$.

(iv) En déduire que si K possède au moins n éléments, alors toute matrice $N \in \mathrm{M}_n(K)$ de trace nulle est un commutateur.

c. On suppose que K est un sous-corps de $\mathbb{C}$. Justifier qu'il existe au moins une matrice diagonale $D \in \mathrm{M}_n(K)$ dont les éléments diagonaux sont deux à deux distincts. Déduire des questions précédentes et de l'exercice 7 qu'une matrice $A \in \mathrm{M}_n(K)$ est de trace nulle si, et seulement si, c'est un commutateur [2].

E9. Soit E un K-espace vectoriel, et soit $u \in \mathscr{L}(E)$.

a. On suppose E de dimension finie. A-t-on toujours $E = \mathrm{Ker}(u) \oplus \mathrm{Im}(u)$?

b. On suppose encore E de dimension finie. Montrer l'équivalence

$$E = \mathrm{Ker}(u) \oplus \mathrm{Im}(u) \iff \mathrm{Ker}(u) = \mathrm{Ker}(u \circ u).$$

c. Le résultat est-il encore vrai si E est de dimension infinie ?

E10. Soit E un K-espace vectoriel, et soit $u \in \mathscr{L}(E)$. Pour tout $k \geqslant 0$, on pose $u^k = u \circ \cdots \circ u$, avec par convention $u^0 = \mathrm{Id}_E$.

a. Montrer que pour tout $k \geqslant 0$, on a

$$\mathrm{Ker}(u^k) \subset \mathrm{Ker}(u^{k+1}) \quad \text{et} \quad \mathrm{Im}(u^{k+1}) \subset \mathrm{Im}(u^k).$$

b. Soit $p \geqslant 0$ un entier vérifiant $\mathrm{Ker}(u^{p+1}) = \mathrm{Ker}(u^p)$. Montrer que pour tout $k \geqslant 0$, on a $\mathrm{Ker}(u^{p+k}) = \mathrm{Ker}(u^p)$.

c. Soit $p \geqslant 0$ vérifiant $\mathrm{Im}(u^{p+1}) = \mathrm{Im}(u^p)$. Montrer que pour tout $k \geqslant 0$, on a $\mathrm{Im}(u^{p+k}) = \mathrm{Im}(u^p)$.

d. Démontrer l'équivalence des propriétés suivantes :

(1) $\mathrm{Ker}(u^2) = \mathrm{Ker}(u)$;

(2) pour tout $k \geqslant 1$, $\mathrm{Ker}(u^k) = \mathrm{Ker}(u)$;

(3) $\mathrm{Ker}(u) \cap \mathrm{Im}(u) = \{0\}$.

e. Démontrer l'équivalence des propriétés suivantes :

(4) $\mathrm{Im}(u^2) = \mathrm{Im}(u)$;

(5) pour tout $k \geqslant 1$, $\mathrm{Im}(u^k) = \mathrm{Im}(u)$;

(6) $E = \mathrm{Ker}(u) + \mathrm{Im}(u)$.

f. Démontrer l'équivalence des propriétés suivantes :

(7) $\mathrm{Ker}(u^2) = \mathrm{Ker}(u)$ et $\mathrm{Im}(u^2) = \mathrm{Im}(u)$;

(8) $E = \mathrm{Ker}(u) \oplus \mathrm{Im}(u)$.

g. On suppose que E est de dimension finie sur K. Montrer que les propriétés (1)-(8) sont toutes équivalentes à :

(9) $\mathrm{rg}(u^2) = \mathrm{rg}(u)$.

---

2. Encore une fois, les arguments présentés sont encore valables si K est de caractéristique nulle.

h. (i) Montrer que $C = \bigcap_{k\geqslant 0} \mathrm{Im}(u^k)$ et $N = \bigcup_{k\geqslant 0} \mathrm{Ker}(u^k)$ sont des sous-espaces de E, stables par u.

(ii) Montrer les équivalences suivantes :

(1) u est injectif $\iff N = \{0\}$.

(2) u est surjectif $\iff C = E$.

i. (i) On suppose qu'il existe un entier $k \geqslant 0$ tel que $\mathrm{Im}(u^k) = \mathrm{Im}(u^{k+1})$. Soit $r(u)$ le plus petit entier vérifiant cette égalité.
Montrer que $u_C \in \mathscr{L}(C)$ est surjectif, et que $E = N + \mathrm{Im}(u^{r(u)})$.

(ii) On suppose qu'il existe $k \geqslant 0$ tel que $\mathrm{Ker}(u^k) = \mathrm{Ker}(u^{k+1})$. On note $s(u)$ le plus petit entier vérifiant cette égalité.
Montrer que $u_N \in \mathscr{L}(N)$ est nilpotent, i.e. qu'il existe $p \geqslant 1$ tel que $u_N^p = 0$, et que l'on a $N \cap \mathrm{Im}(u^{s(u)}) = \{0\}$.

On dit que u est *de caractère fini* s'il existe deux entiers r, $s \geqslant 0$ tels que $\mathrm{Im}(u^r) = \mathrm{Im}(u^{r+1})$ et $\mathrm{Ker}(u^s) = \mathrm{Ker}(u^{s+1})$.

j. (i) On suppose que u est de caractère fini. Montrer que $E = N \oplus C$, que u_N est nilpotent et que u_C est un automorphisme de C.

(ii) Montrer que s'il existe $p \geqslant 0$ tel que

$$\mathrm{Ker}(u^p) = \mathrm{Ker}(u^{p+1}) \quad \text{et} \quad \mathrm{Im}(u^{p+1}) = \mathrm{Im}(u^{p+2}),$$

alors $\mathrm{Im}(u^p) = \mathrm{Im}(u^{p+1})$.

De même, montrer que s'il existe $p \geqslant 0$ tel que

$$\mathrm{Im}(u^p) = \mathrm{Im}(u^{p+1}) \quad \text{et} \quad \mathrm{Ker}(u^{p+1}) = \mathrm{Ker}(u^{p+2}),$$

alors $\mathrm{Ker}(u^p) = \mathrm{Ker}(u^{p+1})$.

(iii) En déduire que si u est de caractère fini, on a $r(u) = s(u)$.

(iv) Montrer que si E est de dimension finie, tout $u \in \mathscr{L}(E)$ est de caractère fini. Si E est de dimension infinie, donner un exemple d'endomorphisme $u \in \mathscr{L}(E)$ tel que $N = C = E$, et un autre tel que $N = C = \{0\}$.

Chapitre IV

Déterminant

Dans ce chapitre, on définit le déterminant d'un endomorphisme ou d'une matrice, en adoptant le point de vue des formes multilinéaires alternées. On étudie alors la structure du groupe linéaire et du groupe spécial linéaire.

On commence par étudier brièvement le groupe symétrique.

1. Introduction éclair au groupe symétrique

Soit $n \geqslant 1$ un entier. On vérifie aisément que l'ensemble $\mathfrak{S}_n$ des bijections de $[\![1, n]\!]$ sur lui-même est un groupe pour la composition des applications.

Pour alléger l'écriture, on notera par Id l'application identité de $[\![1, n]\!]$.

1.1. Définition. Le groupe $\mathfrak{S}_n$ est appelé le *groupe symétrique de degré* n, ou groupe symétrique sur $[\![1, n]\!]$. C'est un groupe à $n!$ éléments.

Un élément $\sigma \in \mathfrak{S}_n$ est appelé une *permutation*.

Une permutation $\tau \in \mathfrak{S}_n$ qui échange deux éléments distincts de $[\![1, n]\!]$, et qui laisse fixe les $n - 2$ autres est appelée une *transposition*.

Si $i \neq j$, on note $(i\ j)$ la transposition qui échange i et j. Remarquons que l'on a $(j\ i) = (i\ j)$, et que $(i\ j) \circ (i\ j) = \mathrm{Id}$, par définition.

Notation. Si $\sigma, \tau \in \mathfrak{S}_n$, on note $\sigma\tau$ l'application composée $\sigma \circ \tau$, et l'on parlera de « produit » de permutations, plutôt que de composition.

On commence par un résultat utile.

1.2. Lemme. *Soit $n \geqslant 1$. Alors, toute permutation $\sigma \in \mathfrak{S}_n$ est un produit d'au plus $n-1$ transpositions.*

Démonstration. On peut toujours supposer que $n \geqslant 2$, le cas $n = 1$ étant clair. Soit $\sigma \in \mathfrak{S}_n$. Supposons que $\sigma \neq \mathrm{Id}$. Il existe donc au moins un élément de $[\![1, n]\!]$ qui ne soit pas fixe par σ. On pose alors

$$k_\sigma = \max\{k \in [\![1, n]\!] \mid \sigma(k) \neq k\}.$$

On pose aussi $k_{\mathrm{Id}} = 0$.

Supposons que $\sigma \neq \mathrm{Id}$, et montrons qu'il existe une transposition $\tau \in \mathfrak{S}_n$ telle que $k_{\tau\sigma} < k_\sigma$. Nous allons montrer que $\tau = \big(k_\sigma\ \sigma(k_\sigma)\big)$ convient. Pour cela, il faut vérifier que $\tau\big(\sigma(j)\big) = j$ pour tout entier $j \geqslant k_\sigma$.

Supposons tout d'abord que $j = k_\sigma$. Alors, σ envoie k_σ sur $\sigma(k_\sigma)$, et τ envoie $\sigma(k_\sigma)$ sur k_σ. Ainsi, on a $\tau\big(\sigma(j)\big) = j$ dans ce cas.

Supposons maintenant que $j > k_\sigma$. Alors, on a $\sigma(j) = j$ par définition même de k_σ. Par choix de j, on a $j \neq k_\sigma$. Si l'on avait $j = \sigma(k_\sigma)$, on aurait $\sigma(j) = j = \sigma(k_\sigma)$, et donc $j = k_\sigma$ par bijectivité de σ, d'où une contradiction. Il en résulte que τ fixe aussi j, d'où $\tau\big(\sigma(j)\big) = j$.

On peut maintenant démontrer le résultat voulu.

Pour $m \in [\![1, n]\!]$, soit (H_m) la proposition suivante.

(H_m) Toute permutation $\sigma \in \mathfrak{S}_n$ telle que $k_\sigma \leqslant m$ est un produit d'au plus $m-1$ transpositions.

On va montrer (H_m) par récurrence (finie). En particulier, (H_n) donnera la conclusion attendue.

Tout d'abord, (H_1) est vraie puisqu'une permutation σ telle que $k_\sigma \leqslant 1$ ne peut être que l'identité, qui est un produit vide de transpositions. En effet, dans ce cas, on a $\sigma(k) = k$ pour tout $k \geqslant 2$ par définition de k_σ. Mais alors, on a nécessairement $\sigma(1) = 1$, car σ est bijective.

Supposons alors que (H_m) soit vraie pour un certain $m \in [\![1, n-1]\!]$, et montrons que (H_{m+1}) est vraie. Soit $\sigma \in \mathfrak{S}_n$ telle que $k_\sigma \leqslant m+1$.

Si $k_\sigma \leqslant m$, alors σ est un produit fini d'au plus $m-1$ transpositions par hypothèse de récurrence, donc d'au plus m transpositions.

Si $k_\sigma = m+1 \geqslant 2$, alors $\sigma \neq \mathrm{Id}$ et il existe une transposition $\tau \in \mathfrak{S}_n$ telle que $k_{\tau\sigma} < k_\sigma$. En particulier, $k_{\tau\sigma} \leqslant m$. Par hypothèse de récurrence, $\tau\sigma$ est un produit d'au plus $m-1$ transpositions. Mais alors, $\sigma = \tau(\tau\sigma)$ est un produit d'au plus m transpositions. Cela achève la récurrence. □

1.3. Remarque. La démonstration précédente donne un algorithme pour décomposer une permutation $\sigma \in \mathfrak{S}_n$ en produit de transpositions.

▷ Si $\sigma = \mathrm{Id}$, on s'arrête.

▷ Sinon, on choisit le plus grand $k \in [\![1, n]\!]$ tel que $\sigma(k) \neq k$.

▷ On remplace σ par $\big(k\ \sigma(k)\big)\sigma$, et l'on recommence.

De proche en proche, on obtient des transpositions $\tau_1, \ldots, \tau_r$ telles que

$$(\tau_r \cdots \tau_1)\sigma = \mathrm{Id}.$$

On a alors $\sigma = \tau_1 \cdots \tau_r$.

La décomposition précédente n'est pas unique, puisque l'on peut toujours rajouter deux copies d'une même transposition, par exemple. En revanche, on a le résultat suivant.

1.4. Proposition. *La parité du nombre de transpositions nécessaires pour décomposer une permutation ne dépend pas de la décomposition choisie.*

Démonstration. Remarquons qu'il suffit de démontrer le résultat suivant : si Id est un produit de p transpositions, alors p est pair.

Supposons en effet ce résultat démontré, et supposons que $\sigma \in \mathfrak{S}_n$ se décompose sous la forme

$$\sigma = \tau_1 \cdots \tau_r = \tau'_1 \cdots \tau'_s,$$

où les τ_i et les τ'_j sont des transpositions. Alors, on a

$$\mathrm{Id} = \tau_1 \cdots \tau_r \tau'_s \cdots \tau'_1.$$

Par le résultat précédent, $r + s$ est pair, i.e. r et s ont même parité.

Soit $n \geqslant 2$. Commençons par quelques considérations générales.

Les lettres i, j et k désignant des éléments distincts de $[\![1, n-1]\!]$, on vérifie les égalités suivantes :

$$(i\ n)(j\ k) = (j\ k)(i\ n)$$
$$(i\ n)(i\ j) = (i\ j)(j\ n)$$
$$(i\ n)(j\ n) = (i\ j)(i\ n)$$
$$(i\ n)(i\ n) = \mathrm{Id}.$$

Bien sûr, certaines de ces égalités n'ont un sens que si $[\![1, n-1]\!]$ a suffisamment d'éléments.

Ces relations permettent de transformer la décomposition d'un élément σ de $\mathfrak{S}_n$ en un produit de p transpositions en « repoussant » les transpositions de la forme $(i\ n)$ vers la droite du produit, en éliminant les produits de deux transpositions de cette forme, et en supprimant les paires de transpositions identiques.

Ainsi, toute permutation $\sigma \in \mathfrak{S}_n$ s'écrivant comme produit de p transpositions se décompose sous la forme

$$\sigma = \tau_1 \cdots \tau_{q-1}\tau_q,$$

où q est un entier de même parité que p, $\tau_1, \ldots, \tau_{q-1}$ sont des transpositions de la forme $(i\ j)$, avec $i, j \in [\![1, n-1]\!]$, et τ_q est une transposition de $\mathfrak{S}_n$.

De plus, σ fixe n si, et seulement si, τ_q fixe n. En effet, si τ_q fixe n, il en est de même de $\sigma = \tau_1 \cdots \tau_{q-1}\tau_q$. Inversement, si σ fixe n, alors la composée $\tau_q = \tau_{q-1} \cdots \tau_1\sigma$ fixe n.

Supposons maintenant que Id soit un produit de p transpositions. Si $p = 0$, on a fini. Sinon, d'après ce qui précède, Id se décompose en produit de q transpositions, avec $q \equiv p\ [2]$, dont au moins les $q-1$ premières fixent n. Comme Id fixe n, la dernière fixe également n. Ainsi, Id se décompose en produit de q transpositions de la forme $(i\ j), i, j \in [\![1, n-1]\!]$, avec $q \equiv p\ [2]$. En itérant ce procédé, on obtient une décomposition en un produit de k copies de la transposition $(1\ 2)$, avec $k \equiv p\ [2]$. Or, le produit de k copies de $(1\ 2)$ donnent Id si k est pair, et $(1\ 2) \neq \mathrm{Id}$ si k est impair. Il en découle que k est pair, et par suite p est pair. Cela achève la démonstration [1]. □

Nous allons maintenant définir la signature d'une permutation.

1.5. Définition. Soit $n \geqslant 1$, et soit $\sigma \in \mathfrak{S}_n$. Si σ se décompose en produit de r transpositions, la *signature* de σ est l'élément

$$\varepsilon(\sigma) = (-1)^r.$$

Par la proposition précédente, cet élément de $\{\pm 1\}$ ne dépend pas de la décomposition choisie.

La proposition suivante résume les propriétés de la signature.

1.6. Proposition. *Soit $n \geqslant 2$. Alors, on a les propriétés suivantes :*

(1) *on a $\varepsilon(\mathrm{Id}) = 1$;*

(2) *toute transposition est de signature -1 ;*

(3) *pour tous $\sigma, \sigma' \in \mathfrak{S}_n$, on a [2] $\varepsilon(\sigma\sigma') = \varepsilon(\sigma)\varepsilon(\sigma')$;*

(4) *pour tout $\sigma \in \mathfrak{S}_n$, on a $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)^{-1}$;*

1. L'exposition très élégante des arguments précédents (qui sont très classiques) est extraite d'une conversation sur le forum les-mathematiques.net : http ://www.les-mathematiques.net/phorum/read.php ?3,515809,516051

2. Dans le langage de la théorie des groupes, on dit que la signature est un morphisme de groupes de $\mathfrak{S}_n$ dans $\{\pm 1\}$.

(5) *l'ensemble* $\mathfrak{A}_n$ *des permutations de signature* 1 *est un sous-groupe*[3] *du groupe* $\mathfrak{S}_n$.

Démonstration. Au vu de la définition, (1) et (2) sont clairs.

Soient $\sigma, \sigma' \in \mathfrak{S}_n$. Si σ et σ' sont produits de r et s transpositions respectivement, alors $\sigma\sigma'$ est produit de $r+s$ transpositions, d'où (3). Les points (1) et (3) impliquent que pour tout $\sigma \in \mathfrak{S}_n$, on a

$$\varepsilon(\sigma)\varepsilon(\sigma^{-1}) = 1.$$

Le point (4) s'en déduit facilement. Il reste à montrer (5).

Le point (1) montre que $\mathrm{Id} \in \mathfrak{A}_n$. De plus, si $\sigma, \sigma' \in \mathfrak{A}_n$, alors $\sigma\sigma' \in \mathfrak{A}_n$ par (3). Enfin, si $\sigma \in \mathfrak{A}_n$, le point (4) montre que $\sigma^{-1} \in \mathfrak{A}_n$.

Cela achève la démonstration. □

1.7. Définition. Le groupe $\mathfrak{A}_n$ des permutations de signature 1 s'appelle le *groupe alterné* de degré n.

On finit ce paragraphe en déterminant le cardinal de $\mathfrak{A}_n$ lorsque $n \geqslant 2$.

1.8. Lemme. *Soit* $n \geqslant 2$, *et soit* $\rho \in \mathfrak{S}_n \setminus \mathfrak{A}_n$. *On note*

$$\mathfrak{A}_n\rho = \{\sigma\rho \mid \sigma \in \mathfrak{A}_n\}.$$

Alors,

$$\mathfrak{S}_n = \mathfrak{A}_n \cup \mathfrak{A}_n\rho,$$

et la réunion est disjointe. En particulier, $\mathfrak{A}_n$ *possède* $\dfrac{n!}{2}$ *éléments.*

Démonstration. Montrons le premier point. Soit $\sigma \in \mathfrak{S}_n$. Si $\sigma \in \mathfrak{A}_n$, on a fini. Si $\sigma \notin \mathfrak{A}_n$, on a $\varepsilon(\sigma) = -1$, et par suite

$$\varepsilon(\sigma\rho^{-1}) = \varepsilon(\sigma)\varepsilon(\rho)^{-1} = (-1)(-1) = 1,$$

par la proposition 1.6. Ainsi, on a $\sigma\rho^{-1} \in \mathfrak{A}_n$, et $\sigma = (\sigma\rho^{-1})\rho \in \mathfrak{A}_n\rho$, d'où l'égalité requise. De plus, l'ensemble $\mathfrak{A}_n \cap \mathfrak{A}_n\rho$ est vide, puisque la signature d'un élément de $\mathfrak{A}_n$ est 1, tandis qu'un élément de $\mathfrak{A}_n\rho$ a une signature égale à -1, comme on le voit par un calcul similaire au précédent.

Puisque ρ est bijective, la multiplication à droite par ρ induit alors une bijection entre $\mathfrak{A}_n$ et $\mathfrak{A}_n\rho$. Le point précédent entraîne alors les égalités

$$\mathrm{card}(\mathfrak{S}_n) = \mathrm{card}(\mathfrak{A}_n) + \mathrm{card}(\mathfrak{A}_n\rho) = 2\mathrm{card}(\mathfrak{A}_n),$$

soit

$$\mathrm{card}(\mathfrak{A}_n) = \frac{\mathrm{card}(\mathfrak{S}_n)}{2} = \frac{n!}{2},$$

d'où le résultat. □

3. En particulier, c'est un groupe pour la composition des applications.

2. Formes multilinéaires alternées, déterminant par rapport à une base

2.1. Définition. Soit E un K-espace vectoriel. Une *forme n-linéaire* sur E est une application

$$\begin{aligned} \varphi : E \times \cdots \times E &\longrightarrow K \\ (x_1, \ldots, x_n) &\longmapsto \varphi(x_1, \ldots, x_n) \end{aligned}$$

qui est linéaire en chaque variable.

Lorsque $n = 1$, on parle de *forme linéaire* sur E. Autrement dit, une forme linéaire est une application linéaire $f : E \longrightarrow K$.

On dit qu'une forme n-linéaire φ sur E est *symétrique* si

$$\varphi(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = \varphi(x_1, \ldots, x_n), \quad \text{pour tous } x_1, \ldots, x_n \in E,\ \sigma \in \mathfrak{S}_n.$$

On dit qu'elle est *antisymétrique* si

$$\varphi(x_{\sigma(1)}, \ldots, x_{\sigma(n)}) = \varepsilon(\sigma)\varphi(x_1, \ldots, x_n)$$

pour tous $x_1, \ldots, x_n \in E, \sigma \in \mathfrak{S}_n$, et *alternée* si

$$\text{pour tous } x_1, \ldots, x_n \in E,\ x_i = x_j \text{ pour } i \neq j \implies \varphi(x_1, \ldots, x_n) = 0.$$

2.2. Remarque. Si φ est alternée, alors elle est antisymétrique. Pour le vérifier, soit $\tau = (ij)$. Alors, on a

$$\begin{aligned} 0 &= \varphi(x_1, \ldots, x_i + x_j, \ldots, x_i + x_j, \ldots, x_n) \\ &= \varphi(x_1, \ldots, x_j, \ldots, x_i, \ldots, x_n) + \varphi(x_1, \ldots, x_i, \ldots, x_j, \ldots, x_n). \end{aligned}$$

La propriété requise est donc vérifiée pour les transpositions. Comme toute permutation de $\mathfrak{S}_n$ est produit de transpositions et que la signature d'un produit est le produit des signatures, on en déduit que la propriété est vraie pour tout $\sigma \in \mathfrak{S}_n$, et que φ est antisymétrique.

Si $2{\cdot}1_K \neq 0_K \in K$, la réciproque est vraie. Si $x_i = x_j = x$, alors on a

$$\varphi(x_1, \ldots, x, \ldots, x, \ldots, x_n) = -\varphi(x_1, \ldots, x, \ldots, x, \ldots, x_n),$$

et donc $2\varphi(x_1, \ldots, x, \ldots, x, \ldots, x_n) = 0$. On multiplie alors par $(2{\cdot}1_K)^{-1}$ pour conclure.

En revanche, si $2{\cdot}1_K = 0_K \in K$, il existe des formes n-linéaires qui sont antisymétriques et non alternées, comme par exemple

$$\begin{aligned} \mathbb{F}_2 \times \mathbb{F}_2 &\longrightarrow \mathbb{F}_2 \\ (\overline{n}, \overline{m}) &\longmapsto \overline{nm}. \end{aligned}$$

2.3. Lemme. *Soient $f_1, ..., f_n$ n formes linéaires sur E. Alors, l'application*

$$\begin{aligned}\varphi_{f_1,\dots,f_n} : E \times \cdots \times E &\longrightarrow K \\ (x_1, \dots, x_n) &\longmapsto \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) f_{\sigma(1)}(x_1) \cdots f_{\sigma(n)}(x_n)\end{aligned}$$

est n-linéaire alternée.

Démonstration. Pour alléger les notations, on écrira φ à la place de $\varphi_{f_1,\dots,f_n}$. La n-linéarité de φ est facile et laissée au lecteur. Supposons maintenant que l'on ait $x_i = x_j$ pour $i < j$. Soit $\tau = (ij)$. Puisque $\mathfrak{S}_n$ est la réunion disjointe de $\mathfrak{A}_n$ et de $\mathfrak{A}_n\tau$, par le lemme 1.8, $\varphi(x_1, \dots, x_n)$ est égal à

$$\sum_{\sigma \in \mathfrak{A}_n} \big(\varepsilon(\sigma) f_{\sigma(1)}(x_1) \cdots f_{\sigma(n)}(x_n) + \varepsilon(\sigma\tau) f_{\sigma(\tau(1))}(x_1) \cdots f_{\sigma(\tau(n))}(x_n)\big).$$

Comme $\sigma \in \mathfrak{A}_n$, on a $\varepsilon(\sigma) = 1$ et $\varepsilon(\sigma\tau) = \varepsilon(\sigma)\varepsilon(\tau) = -1$. On a ainsi

$$\varphi(x_1, \dots, x_n) = \sum_{\sigma \in \mathfrak{A}_n} \big(f_{\sigma(1)}(x_1) \cdots f_{\sigma(n)}(x_n) - f_{\sigma(\tau(1))}(x_1) \cdots f_{\sigma(\tau(n))}(x_n)\big).$$

Remarquons maintenant que $f_{\sigma(\tau(k))}(x_k) = f_{\sigma(k)}(x_k)$ si $k \neq i, j$. De plus, on a

$$f_{\sigma(\tau(i))}(x_i) = f_{\sigma(j)}(x_i) = f_{\sigma(j)}(x_j),$$

car $x_i = x_j$, et de même

$$f_{\sigma(\tau(j))}(x_j) = f_{\sigma(i)}(x_j) = f_{\sigma(i)}(x_i).$$

Ainsi, on a

$$f_{\sigma(1)}(x_1) \cdots f_{\sigma(n)}(x_n) = f_{\sigma(\tau(1))}(x_1) \cdots f_{\sigma(\tau(n))}(x_n),$$

si bien que l'on obtient finalement

$$\varphi(x_1, \dots, x_n) = 0.$$

Cela achève la démonstration. □

2.4. Théorème. *Soit E un K-espace vectoriel de dimension $n \geqslant 1$. Alors, le K-espace vectoriel $\mathrm{Alt}_n(E)$ des formes n-linéaires alternées est de dimension 1.*

Plus précisément, pour tout base $\mathbf{e} = (e_1, \dots, e_n)$ de E, il existe une unique forme n-linéaire alternée $\det_{\mathbf{e}}$ telle que

$$\det_{\mathbf{e}}(e_1, \dots, e_n) = 1,$$

et elle engendre $\mathrm{Alt}_n(E)$. Si $x_j = \sum_{i=1}^{n} a_{ij} e_i$, on a

$$\det_{\mathbf{e}}(x_1, \dots, x_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}.$$

Démonstration. Montrons tout d'abord que $\det_{\mathbf{e}}$ est une forme n-linéaire alternée vérifiant

$$\det_{\mathbf{e}}(e_1, \ldots, e_n) = 1.$$

On a en fait

$$\det_{\mathbf{e}} = \varphi_{e_1^*, \ldots, e_n^*},$$

où $\mathbf{e}^* = (e_1, \ldots, e_n)$ est la base duale de $\mathbf{e}$. En effet, soient $x_1, \ldots, x_n \in E$. Écrivons $x_j = \sum_{i=1}^{n} a_{ij} e_i$. Alors, on a $e^*_{\sigma(k)}(x_k) = a_{\sigma(k)k}$, d'où l'égalité annoncée. Par le lemme précédent, on obtient que $\det_{\mathbf{e}}$ est n-linéaire alternée. De plus, on a

$$\det_{\mathbf{e}}(e_1, \ldots, e_n) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) \delta_{\sigma(1)1} \cdots \delta_{\sigma(n)n}.$$

Si σ a un point qui n'est pas fixe, alors le terme correspondant est nul. Ainsi, le seul terme de la somme qui contribue non trivialement est celui qui correspond à $\sigma = \mathrm{Id}$, et celui ci vaut 1.

Soit maintenant φ une forme n-linéaire alternée. Puisque φ est n-linéaire, on a

$$\varphi\Big(\sum_{i=1}^{n} a_{i1} e_i, \ldots, \sum_{i=1}^{n} a_{in} e_i\Big) = \sum_{i_1, \ldots, i_n} a_{i_1 1} \cdots a_{i_n n} \varphi(e_{i_1}, \ldots, e_{i_n}).$$

Puisque φ est alternée, tous ces termes sont nuls, sauf éventuellement si les indices $i_1, \ldots, i_n$ sont tous distincts, autrement dit, sauf s'il existe une permutation $\sigma \in \mathfrak{S}_n$ telle que $i_j = \sigma(j)$ pour tout $j \in [\![1, n]\!]$. On a ainsi

$$\varphi\Big(\sum_{i=1}^{n} a_{i1} e_i, \ldots, \sum_{i=1}^{n} a_{in} e_i\Big) = \sum_{\sigma \in \mathfrak{S}_n} a_{\sigma(1)1} \cdots a_{\sigma(n)n} \varphi(e_{\sigma(1)}, \ldots, e_{\sigma(n)}).$$

Comme φ est alternée, elle est aussi antisymétrique, et l'on obtient donc

$$\varphi\Big(\sum_{i=1}^{n} a_{i1} e_i, \ldots, \sum_{i=1}^{n} a_{in} e_i\Big) = \varphi(e_1, \ldots, e_n) \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}.$$

Par conséquent, on a $\varphi = \varphi(e_1, \ldots, e_n) \det_{\mathbf{e}}$. Ainsi, le K-espace vectoriel des formes n-linéaires alternées est de dimension 1, engendré par $\det_{\mathbf{e}}$. L'égalité précédente montre alors que si φ vérifie $\varphi(e_1, \ldots, e_n) = 1$, alors $\varphi = \det_{\mathbf{e}}$. Cela achève la démonstration. □

2.5. Définition. Soit E un K-espace de dimension n, soit $\mathbf{e}$ une base de E, et soient $x_1, \ldots, x_n \in E$. Alors, $\det_{\mathbf{e}}(x_1, \ldots, x_n)$ s'appelle le *déterminant* des vecteurs $x_1, \ldots, x_n \in E$ par rapport à la base $\mathbf{e}$.

2.6. Lemme. *Soit E un K-espace vectoriel. Pour toutes bases $\mathbf{e}$, $\mathbf{e}'$ de E,*

on a

$$\det_{\mathbf{e}'} = \det_{\mathbf{e}'}(e_1, \ldots, e_n) \det_{\mathbf{e}}.$$

Démonstration. Soient $\mathbf{e}$, $\mathbf{e}'$ deux bases de E. Puisque $\det_{\mathbf{e}} \in \mathrm{Alt}_n(E)$ est une base de $\mathrm{Alt}_n(E)$ par le théorème 2.4, il existe $\lambda \in K$ tel que

$$\det_{\mathbf{e}'} = \lambda \det_{\mathbf{e}}.$$

Il suffit alors d'appliquer cette égalité aux vecteurs $e_1, \ldots, e_n$ pour en déduire le résultat, puisque $\det_{\mathbf{e}}(e_1, \ldots, e_n) = 1$. □

2.7. Corollaire. *Soit E un K-espace de dimension $n \geqslant 1$, soit $\mathbf{e}$ une base de E, et soient $x_1, \ldots, x_n \in E$. Alors, on a*

$$\begin{aligned} \det_{\mathbf{e}}(x_1, \ldots, x_n) \neq 0 &\iff (x_1, \ldots, x_n) \text{ est libre} \\ &\iff (x_1, \ldots, x_n) \text{ est une base de } E. \end{aligned}$$

Démonstration. La deuxième équivalence est bien connue. Supposons que les vecteurs $x_1, \ldots, x_n$ soient liés. Alors, il existe $i \in [\![1, n]\!]$ tel que

$$x_i = \sum_{j \neq i} \lambda_j x_j, \quad \lambda_j \in K.$$

On a alors

$$\det_{\mathbf{e}}(x_1, \ldots, x_n) = \sum_{j \neq i} \lambda_j \det_{\mathbf{e}}(x_1, \ldots, x_j, \ldots, x_j, \ldots, x_n) = 0,$$

car $\det_{\mathbf{e}}$ est n-linéaire alternée.

Supposons maintenant que $\mathbf{x} = (x_1, \ldots, x_n)$ soit une famille libre. C'est donc une base puisque $\dim_K(E) = n$. D'après le lemme précédent, on a

$$\det_{\mathbf{e}} = \det_{\mathbf{e}}(x_1, \ldots, x_n) \det_{\mathbf{x}}.$$

Ainsi, on a

$$1 = \det_{\mathbf{e}}(x_1, \ldots, x_n) \det_{\mathbf{x}}(e_1, \ldots, e_n),$$

ce qui implique en particulier que $\det_{\mathbf{e}}(x_1, \ldots, x_n) \neq 0$. □

3. Déterminant d'un endomorphisme

Soit E un K-espace vectoriel de dimension n, et soit $\mathbf{e}$ une base de E. Pour tout $u \in \mathscr{L}(E)$, l'application

$$\begin{aligned} E \times \cdots \times E &\longrightarrow K \\ (x_1, \ldots, x_n) &\longmapsto \det_{\mathbf{e}}\big(u(x_1), \ldots, u(x_n)\big) \end{aligned}$$

est n-linéaire alternée.

Il existe donc un **unique** élément $\det_{\mathbf{e}}(u) \in K$ tel que

$$\det_{\mathbf{e}}\big(u(x_1),\dots,u(x_n)\big) = \det_{\mathbf{e}}(u)\det_{\mathbf{e}}(x_1,\dots,x_n)$$

pour tous $x_1,\dots,x_n \in E$. Autrement dit, on a

$$\det_{\mathbf{e}}(u) = \det_{\mathbf{e}}\big(u(e_1),\dots,u(e_n)\big).$$

Il suffit en effet d'appliquer l'égalité précédente à $(e_1,\dots,e_n)$ pour le voir.

3.1. Lemme. *Soit $u \in \mathscr{L}(E)$ et soit $\mathbf{e}$ une base de E. Alors, $\det_{\mathbf{e}}(u)$ est indépendant du choix de $\mathbf{e}$.*

Démonstration. Soit $\mathbf{e}'$ une autre base de E. Par définition de $\det_{\mathbf{e}}(u)$, on a l'égalité

$$\det_{\mathbf{e}}\big(u(e'_1),\dots,u(e'_n)\big) = \det_{\mathbf{e}}(u)\det_{\mathbf{e}}(e'_1,\dots,e'_n).$$

Par le lemme 2.6 appliqué à $(e'_1,\dots,e'_n)$, on a

$$\det_{\mathbf{e}}(e'_1,\dots,e'_n)\det_{\mathbf{e}'}(e_1,\dots,e_n) = 1.$$

L'égalité précédente se récrit alors

$$\det_{\mathbf{e}}(u) = \det_{\mathbf{e}'}(e_1,\dots,e_n)\det_{\mathbf{e}}\big(u(e'_1),\dots,u(e'_n)\big).$$

Mais par le même lemme appliqué cette fois à $\big(u(e'_1),\dots,u(e'_n)\big)$, on en déduit que

$$\det_{\mathbf{e}}(u) = \det_{\mathbf{e}'}\big(u(e'_1),\dots,u(e'_n)\big) = \det_{\mathbf{e}'}(u).$$

Cela achève la démonstration. □

3.2. Définition. Soit E un K-espace vectoriel de dimension n. Le *déterminant* d'un endomorphisme $u \in \mathscr{L}(E)$, noté $\det(u)$, est l'élément de K défini par

$$\det(u) = \det_{\mathbf{e}}(u) = \det_{\mathbf{e}}\big(u(e_1),\dots,u(e_n)\big),$$

où $\mathbf{e} = (e_1,\dots,e_n)$ est une base arbitraire de E.

3.3. Proposition. *Soit E un K-espace vectoriel de dimension finie n, soient $u, v \in \mathscr{L}(E)$ et soit $\lambda \in K$. Alors, on a les propriétés suivantes :*

(1) $\det(\mathrm{Id}_E) = 1$;

(2) $\det(u \circ v) = \det(u)\det(v) = \det(v \circ u)$;

(3) $\det(\lambda u) = \lambda^n \det(u)$;

(4) *on a $\det(u) \neq 0$ si, et seulement si, $u \in \mathrm{GL}(E)$. Dans ce cas,*

$$\det(u^{-1}) = \det(u)^{-1}.$$

Démonstration. Soit $\mathbf{e}$ une base de E. Le premier point est clair, et découle de la définition.

Pour montrer (2), il suffit de remarquer que pour tous vecteurs $x_1, \ldots, x_n$ dans E, on a

$$\begin{aligned} \det_{\mathbf{e}}\big((u\circ v)(x_1),\ldots,(u\circ v)(x_n)\big) &= \det_{\mathbf{e}}\big(u(v(x_1)),\ldots,u(v(x_n))\big) \\ &= \det_{\mathbf{e}}(u)\det_{\mathbf{e}}\big(v(x_1),\ldots,v(x_n)\big) \\ &= \det_{\mathbf{e}}(u)\det_{\mathbf{e}}(v)\det_{\mathbf{e}}(x_1,\ldots,x_n). \end{aligned}$$

On en déduit ainsi (2) en appliquant la définition de $\det_{\mathbf{e}}(u\circ v)$. Montrons le point (3). Pour tous $x_1, \ldots, x_n \in E$, on a

$$\begin{aligned} \det_{\mathbf{e}}\big(\lambda u(x_1),\ldots,\lambda u(x_n)\big) &= \lambda^n \det_{\mathbf{e}}\big(u(x_1),\ldots,u(x_n)\big) \\ &= \lambda^n \det_{\mathbf{e}}(u)\det_{\mathbf{e}}(x_1,\ldots,x_n). \end{aligned}$$

On en déduit le résultat.

Montrons (4). Par le corollaire 2.7, on a $\det_{\mathbf{e}}(u) \neq 0$ si, et seulement si, la famille $\big(u(e_1),\ldots,u(e_n)\big)$ est une base de E. Cela équivaut au fait que u soit inversible par le théorème III-2.4. L'égalité manquante découle alors des propriétés (1) et (2). □

On peut maintenant définir le groupe spécial linéaire.

3.4. Définition. Le *groupe spécial linéaire* de E, noté $\mathrm{SL}(E)$, est l'ensemble des automorphismes de E de déterminant 1. On montre aisément que c'est un sous-groupe[4] de $\mathrm{GL}(E)$.

4. Déterminant d'une matrice

4.1. Définition. Soit $M \in \mathrm{M}_n(K)$. Le *déterminant* de M est le déterminant de ses vecteurs colonnes par rapport à la base canonique de K^n. Il est noté $\det(M)$. C'est aussi le déterminant de l'endomorphisme

$$\begin{aligned} u : K^n &\longrightarrow K^n \\ X &\longmapsto MX. \end{aligned}$$

Autrement dit, si $M = (a_{ij})$, on a

$$\det(M) = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1)1}\cdots a_{\sigma(n)n}.$$

Remarquons que, si E est un K-espace vectoriel de dimension finie, pour tout $u \in \mathscr{L}(E)$, on a

$$\det(u) = \det\big(\mathrm{Mat}(u\,;\mathbf{e})\big)$$

pour toute base $\mathbf{e}$ de E.

4. Cela se démontre en utilisant la proposition précédente, et des arguments similaires à ceux utilisés pour démontrer le dernier point de la proposition 1.6.

En particulier, $u \in \mathrm{SL}(E)$ si, et seulement si, $\mathrm{Mat}(u\,;\mathbf{e})$ est une matrice de déterminant 1.

On note $\mathrm{GL}_n(K)$ l'ensemble des matrices inversibles de $\mathrm{M}_n(K)$, et $\mathrm{SL}_n(K)$ l'ensemble des matrices de $\mathrm{M}_n(K)$ de déterminant 1. Ce sont des groupes pour la multiplication matricielle.

On dénotera dans la suite aussi par « det » l'application déterminant par rapport à la base canonique de K^n.

On a alors les propriétés suivantes.

4.2. Proposition. *Soit K un corps et soit $n \geqslant 1$ un entier. Pour toutes matrices M, $N \in \mathrm{M}_n(K)$ et tout $\lambda \in K$, on a*

(1) $\det(\mathrm{I}_n) = 1$;

(2) $\det(MN) = \det(M)\det(N) = \det(NM)$;

(3) $\det(\lambda M) = \lambda^n \det(M)$;

(4) *on a* $\det(M) \neq 0$ *si, et seulement si, M est inversible. Dans ce cas,*

$$\det(M^{-1}) = \det(M)^{-1}$$

(5) $\det(M^t) = \det(M)$;

(6) *si deux colonnes/lignes de M sont identiques, alors* $\det(M) = 0$;

(7) *si les colonnes/lignes de M sont liées, alors* $\det(M) = 0$;

(8) *lorsque l'on échange deux colonnes/lignes de M, on multiplie la valeur du déterminant par -1 ;*

(9) *lorsque l'on multiplie une colonne/ligne par $\lambda \in K$, on multiplie par λ la valeur du déterminant ;*

(10) *si $C_1, \ldots, C_n$ sont les colonnes de M, et si $L_1, \ldots, L_n$ sont les lignes de M, alors pour tous $i, j \in [\![1, n]\!]$, où $i \neq j$, et pour tout $\lambda \in K$, les opérations*

$$C_i \longleftarrow C_i + \lambda C_j, \quad L_i \longleftarrow L_i + \lambda L_j$$

ne changent pas la valeur du déterminant.

Démonstration. Les points (1) – (4) sont une retraduction matricielle de la proposition 3.3.

Montrons le point (5). Écrivons $M = (a_{ij})$, et donc $M^t = (a_{ji})$. Si $\sigma \in \mathfrak{S}_n$, alors σ^{-1} induit une bijection de $[\![1, n]\!]$ sur lui-même. On a donc

$$\det(M) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n} = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{1\sigma^{-1}(1)} \cdots a_{n\sigma^{-1}(n)}.$$

Maintenant, l'application

$$\begin{aligned} \mathfrak{S}_n &\longrightarrow \mathfrak{S}_n \\ \sigma &\longmapsto \sigma^{-1} \end{aligned}$$

est bijective, et l'on a donc

$$\sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)a_{\sigma(1)1}\cdots a_{\sigma(n)n} = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma^{-1})a_{1\sigma(1)}\cdots a_{n\sigma(n)}.$$

Par ailleurs, on a $\varepsilon(\sigma^{-1}) = \varepsilon(\sigma)^{-1} = \varepsilon(\sigma)$, la dernière égalité provenant du fait que la signature prend ses valeurs dans l'ensemble $\{\pm 1\}$. On obtient alors

$$\det(M) = \sum_{\sigma\in\mathfrak{S}_n} \varepsilon(\sigma)a_{1\sigma(1)}\cdots a_{n\sigma(n)} = \det(M^t).$$

Les points (6) – (10) proviennent du fait que le déterminant est une application n-linéaire alternée en ce qui concerne les propriétés sur les colonnes, qui sont aussi alors vraies sur les lignes en passant à la transposée en vertu du point (5). □

4.3. Exemple. Pour tous a, b, c, $d \in K$, on a

$$\det(a) = a \quad \text{et} \quad \det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Les propriétés précédentes sont fort utiles pour calculer le déterminant d'une matrice sur les lignes et les colonnes. Nous allons en donner une autre, qui permet de réduire le calcul du déterminant d'une matrice à celui de matrices de tailles inférieures.

On a tout d'abord besoin d'une définition.

4.4. Définition. Soit $M \in \mathrm{M}_n(K)$. Pour tous i, $j \in [\![1, n]\!]$, on note $\Delta_{ij}(M)$ le déterminant de la sous-matrice de M obtenue en enlevant la ligne i et la colonne j. L'élément $\Delta_{ij}(M)$ s'appelle un *mineur* de M.

4.5. Proposition. *Soit $M = (a_{ij}) \in \mathrm{M}_n(K)$. Alors,*

(1) *pour tout $j \in [\![1, n]\!]$, on a* $\det(M) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij}\Delta_{ij}(M)$ (développement par rapport à la colonne j) ;

(2) *pour tout $i \in [\![1, n]\!]$, on a* $\det(M) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij}\Delta_{ij}(M)$ (développement par rapport à la ligne i).

Démonstration. Démontrons tout d'abord (1). Soient $C_1, \ldots, C_n$ les colonnes de M. On note $(\varepsilon_1, \ldots, \varepsilon_n)$ la base canonique de K^n. Puisque le déterminant est n-linéaire et $C_j = \sum_{i=1}^{n} a_{ij}\varepsilon_i$, on a

$$\det(M) = \sum_{i=1}^{n} a_{ij} \det(C_1, \ldots, C_{j-1}, \varepsilon_i, C_{j+1}, \ldots, C_n).$$

Échanger deux colonnes multiplie le déterminant par -1. En faisant $j-1$ échanges de colonnes successifs, on obtient donc

$$\begin{aligned} \det(M) &= \sum_{i=1}^{n} a_{ij}(-1)^{j-1} \det(\varepsilon_i, C_1, \ldots, C_{j-1}, C_{j+1}, \ldots C_n) \\ &= \sum_{i=1}^{n} a_{ij}(-1)^{j-1} \det(P_i), \end{aligned}$$

où $P_i \in \mathrm{M}_n(K)$ est la matrice dont les colonnes sont

$$\varepsilon_i, C_1, \ldots, C_{j-1}, C_{j+1}, \ldots, C_n.$$

En faisant cette fois $i-1$ échanges de lignes successifs dans la matrice P_i, on obtient une matrice de la forme

$$\begin{pmatrix} 1 & * & \cdots & * \\ 0 & & & \\ \vdots & & M_{ij} & \\ 0 & & & \end{pmatrix},$$

où M_{ij} est la matrice obtenue de M en ôtant la ligne i et la colonne j. Comme échanger deux lignes multiplie le déterminant par -1, on a donc

$$\det(P_i) = (-1)^{i-1} \det \begin{pmatrix} 1 & * & \cdots & * \\ 0 & & & \\ \vdots & & M_{ij} & \\ 0 & & & \end{pmatrix}.$$

Il reste donc à montrer l'égalité

$$\det \begin{pmatrix} 1 & * & \cdots & * \\ 0 & & & \\ \vdots & & M_{ij} & \\ 0 & & & \end{pmatrix} = \det(M_{ij}).$$

En faisant des opérations élémentaires sur les colonnes, ce qui ne change

pas le déterminant, on obtient

$$\det\begin{pmatrix} 1 & * & \cdots & * \\ 0 & & & \\ \vdots & & M_{ij} & \\ 0 & & & \end{pmatrix} = \det\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & M_{ij} & \\ 0 & & & \end{pmatrix}.$$

Il suffit donc de montrer que pour toute matrice $N \in \mathrm{M}_{n-1}(K)$, on a

$$\det\begin{pmatrix} 1 & \\ & N \end{pmatrix} = \det(N).$$

Or, l'application

$$\begin{aligned} \varphi : K^{n-1} \times \cdots \times K^{n-1} &\longrightarrow K \\ (Y_1, \ldots, Y_{n-1}) &\longmapsto \det\begin{pmatrix} 1 & \\ & Y \end{pmatrix}, \end{aligned}$$

où Y est la matrice dont les colonnes sont $Y_1, \ldots, Y_{n-1}$, est $(n-1)$-linéaire alternée, et il existe donc $\lambda \in K$ tel que

$$\varphi(Y_1, \ldots, Y_{n-1}) = \lambda \det(Y_1, \ldots, Y_{n-1}) = \lambda \det(Y)$$

pour tous $Y_1, \ldots, Y_{n-1} \in K^n$. En prenant $Y = \mathrm{I}_{n-1}$, on obtient

$$\lambda = \det(\mathrm{I}_n) = 1.$$

On a alors le résultat voulu.

Pour montrer (2), on applique (1) à M^t. Cela achève la démonstration. □

4.6. Définition. Si $M \in \mathrm{M}_n(K)$, la *comatrice* de M est la matrice

$$\mathrm{com}(M) = \big((-1)^{i+j}\Delta_{ij}(M)\big) \in \mathrm{M}_n(K).$$

4.7. Proposition. *Pour tout $M \in \mathrm{M}_n(K)$, on a*

$$M\,\mathrm{com}(M)^t = \mathrm{com}(M)^t M = \det(M)\mathrm{I}_n.$$

En particulier, si M est inversible, on a

$$M^{-1} = \big(\det(M)\big)^{-1}\mathrm{com}(M)^t.$$

Démonstration. Le terme général c_{ij} de la matrice $\mathrm{com}(M)^t M$ est

$$c_{ij} = \sum_{k=1}^{n} (-1)^{k+i}\Delta_{ki}(M)a_{kj}.$$

La proposition précédente montre que $c_{ii} = \det(M)$ pour tout $i \in [\![1, n]\!]$. Supposons maintenant que $i \neq j$, et considérons $M' = (b_{ij})$ la matrice obtenue de M en remplaçant la colonne i par la colonne j.

Puisque M' possède deux colonnes identiques, on a $\det(M') = 0$. Mais en développant ce déterminant par rapport à la colonne i, on obtient

$$\det(M') = 0 = \sum_{k=1}^{n} (-1)^{k+i} b_{ki} \Delta_{ki}(M').$$

Or, par construction de M', on a $\Delta_{ki}(M') = \Delta_{ki}(M)$ et $b_{ki} = a_{kj}$ pour tout $k \in [\![1, n]\!]$. Ainsi, on obtient

$$0 = \sum_{k=1}^{n} (-1)^{k+j} a_{kj} \Delta_{ki}(M) = c_{ij}.$$

Ainsi, on obtient $c_{ij} = \det(M)\delta_{ij}$, et donc

$$\mathrm{com}(M)^t M = \det(M) \mathrm{I}_n.$$

L'autre égalité se montre de la manière analogue. Le dernier point est clair, puisque M est inversible si, et seulement si, $\det(M) \neq 0$. □

Nous continuons ce paragraphe en calculant le déterminant d'une matrice triangulaire par blocs.

4.8. Proposition. *Soit $M = (M_{ij})_{1 \leqslant i,j \leqslant n}$ une matrice triangulaire supérieure (resp. inférieure) par blocs, c'est-à-dire $M_{ij} = 0$ pour tout $i > j$ (resp. $i < j$). Alors, on a*

$$\det(M) = \det(M_{11}) \cdots \det(M_{nn}).$$

En particulier, si $M = (a_{ij})$ est une matrice triangulaire, on a

$$\det(M) = a_{11} \cdots a_{nn}.$$

Démonstration. Il suffit de le montrer pour les matrices triangulaires supérieures, l'autre cas s'obtenant en transposant. Il suffit aussi de traiter le cas

$$M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix},$$

le cas général s'obtenant par récurrence.

Soit p la taille de la matrice A. Considérons d'abord le cas $A = \mathrm{I}_p$. Alors, en développant successivement par rapport aux colonnes $1, 2, \ldots, p$, on obtient que $\det(M) = \det(C)$.

Passons au cas général. Si $X_1, \ldots, X_p \in K^p$, soit $X \in \mathrm{M}_p(K)$ la matrice dont les colonnes sont $X_1, \ldots, X_p$.

L'application

$$\begin{aligned} \varphi : K^p \times \cdots \times K^p &\longrightarrow K \\ (X_1, \ldots, X_p) &\longmapsto \det \begin{pmatrix} X & B \\ 0 & C \end{pmatrix} \end{aligned}$$

est p-linéaire alternée, et l'on a donc

$$\varphi = \varphi(\varepsilon_1, \ldots, \varepsilon_p) \det = \det(C) \det,$$

la dernière égalité provenant du point précédent. En appliquant cela aux colonnes de A, on obtient

$$\det(M) = \det(C) \det(A),$$

ce qu'il fallait vérifier. Le dernier point est alors clair. □

4.9. Remarque. Pour calculer le déterminant d'une matrice M, on a intérêt à effectuer des opérations sur les lignes et les colonnes et ainsi faire apparaître un maximum de 0 afin, soit de rendre la matrice M triangulaire (grâce à la méthode du pivot de Gauss), soit de développer par rapport à une ligne ou une colonne bien choisie pour se ramener au calcul du déterminant d'une matrice de plus petite taille.

5. Transvections et dilatations

Nous allons étudier ici brièvement quelques éléments particuliers de $\mathrm{GL}(E)$ et $\mathrm{SL}(E)$. On commence par définir les transvections.

5.1. Définition. Soit E un K-espace vectoriel non nul. Soit H un hyperplan de E, et soit D une droite de E telle que $D \subset H$ (ce qui impose que $\dim_K(E) \geqslant 2$). On dit que $u \in \mathscr{L}(E)$ est **une** *transvection* d'hyperplan H et de droite D si les conditions suivantes sont vérifiées :

(1) $u \neq \mathrm{Id}_E$;

(2) on a $u(x) = x$, pour tout $x \in H$;

(3) on a $u(x) - x \in D$, pour tout $x \in E$.

On dit que $u \in \mathscr{L}(E)$ est une transvection s'il existe un hyperplan H et une droite D de E incluse dans H tels que u soit une transvection d'hyperplan H et de droite D.

On va maintenant caractériser les transvections de manière un peu plus explicite.

5.2. Proposition. *Soit E un K-espace vectoriel non nul. Pour tout endomorphisme $u \in \mathscr{L}(E)$, les propriétés suivantes sont équivalentes :*

(1) *l'endomorphisme u est une transvection ;*

(2) *le sous-espace $\mathrm{Ker}(u-\mathrm{Id}_E)$ est un hyperplan et $(u-\mathrm{Id}_E)\circ(u-\mathrm{Id}_E) = 0$. Dans ce cas, u est une transvection d'hyperplan $\mathrm{Ker}(u - \mathrm{Id}_E)$ et de droite $\mathrm{Im}(u - \mathrm{Id}_E)$;*

(3) *il existe une forme linéaire non nulle $\varphi : E \longrightarrow K$ et $a \in \mathrm{Ker}(\varphi)$ non nul tels que*
$$u(x) = x + \varphi(x)a, \qquad \textit{pour tout } x \in E.$$
Dans ce cas, u est une transvection d'hyperplan $\mathrm{Ker}(\varphi)$ et de droite Ka.

Si de plus E est de dimension finie, ces propriétés sont aussi équivalentes aux propriétés suivantes :

(4) *il existe une base de E dans laquelle la matrice représentative de u est de la forme*
$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & 1 \\ & & & 1 \end{pmatrix},$$
où les coefficients non diagonaux n'apparaissant pas sont nuls ;

(5) *il existe une base de E dans laquelle la matrice représentative de u est de la forme*
$$T_{i,j}(\lambda) = \begin{array}{cc} & \begin{array}{ccc} & & j \\ & & \downarrow \end{array} \\ i \rightarrow & \begin{pmatrix} 1 & & & \\ & \ddots & & \lambda \\ & & \ddots & \\ & & & 1 \end{pmatrix} \end{array},$$
où $i \neq j$ et $\lambda \in K^\times$.

Démonstration

(1) $\Longrightarrow$ (2). Puisque u est une transvection, il existe un hyperplan H et une droite D incluse dans H tels que $u(x) = x$ pour tout $x \in H$ et $u(x) - x \in D$ pour tout $x \in E$. On a donc $H \subset \mathrm{Ker}(u - \mathrm{Id}_E)$ et $\mathrm{Im}(u - \mathrm{Id}_E) \subset D$.

Puisque H est un hyperplan, il existe $x_0 \in E \setminus \{0\}$ tel que $E = H \oplus Kx_0$. Si l'on avait $u(x_0) = x_0$, alors on aurait aisément $u = \mathrm{Id}_E$, ce qui n'est pas le cas. Donc, $u(x_0) - x_0 \neq 0$. Soit $x \in E$, et écrivons
$$x = x_H + \mu x_0, \qquad \text{avec } x_H \in H \text{ et } \mu \in K.$$
On a $u(x) - x = \mu(u(x_0) - x_0)$, et par conséquent $\mathrm{Im}(u - \mathrm{Id}_E)$ est la droite de E engendrée par $u(x_0) - x_0$.

Mais puisque $\mathrm{Im}(u-\mathrm{Id}_E) \subset D$ et que $\dim_K\big(\mathrm{Im}(u-\mathrm{Id}_E)\big) = \dim_K(D) = 1$, on en déduit que $D = \mathrm{Im}(u-\mathrm{Id}_E)$.

Par ailleurs, on a

$$u(x) - x = 0 \iff \mu(u(x_0) - x_0) = 0 \iff \mu = 0 \iff x \in H.$$

Ainsi, $H = \mathrm{Ker}(u-\mathrm{Id}_E)$. En particulier, $\mathrm{Ker}(u-\mathrm{Id}_E)$ est un hyperplan. De plus, puisque $D = \mathrm{Im}(u-\mathrm{Id}_E) \subset H = \mathrm{Ker}(u-\mathrm{Id}_E)$, on en déduit aisément que $(u-\mathrm{Id}_E)\circ(u-\mathrm{Id}_E) = 0$.

(2) $\Longrightarrow$ (3). Puisque $H = \mathrm{Ker}(u-\mathrm{Id}_E)$ est un hyperplan, il existe une forme linéaire non nulle $\varphi : E \longrightarrow K$ telle que $\mathrm{Ker}(\varphi) = H$, et cela par le lemme III-5.8. On a alors $E = H \oplus Kx_0$, où $x_0 \in E$ vérifie $\varphi(x_0) = 1$ par ce même lemme. Posons $a = u(x_0) - x_0$.

Puisque $(u-\mathrm{Id}_E)\circ(u-\mathrm{Id}_E) = 0$, on a $a \in \mathrm{Ker}(u-\mathrm{Id}_E) = \mathrm{Ker}(\varphi)$.

Nous allons montrer que φ et a conviennent. Pour tout $x = x_H + \mu x_0 \in E$, on a $\varphi(x) = \mu$, et donc

$$x + \varphi(x)a = (x_H + \mu x_0) + \mu(u(x_0) - x_0) = x_H + \mu u(x_0).$$

Mais par ailleurs, $H = \mathrm{Ker}(u-\mathrm{Id}_E)$, et nous avons donc

$$u(x) = u(x_H + \mu x_0) = x_H + \mu u(x_0) = x + \varphi(x)a,$$

d'où le résultat.

(3) $\Longrightarrow$ (1). Supposons que

$$u(x) = x + \varphi(x)a, \quad \text{pour tout } x \in E,$$

où $\varphi : E \longrightarrow K$ est une forme linéaire non nulle, et $a \in \mathrm{Ker}(\varphi)$ est non nul. Puisque φ est non nulle et a est non nul, $u \neq \mathrm{Id}_E$.

Posons $D = Ka$ et $H = \mathrm{Ker}(\varphi)$. Par le lemme III-5.8, H est un hyperplan. De plus, puisque $a \in H \setminus \{0\}$, D est une droite de E incluse dans H. Enfin, pour tout $x \in E$, on a $u(x) - x \in Ka = D$, et pour tout $x \in H$, on a $u(x) = x$, puisque $H = \mathrm{Ker}(\varphi)$. Ainsi, u est bien une transvection d'hyperplan $\mathrm{Ker}(\varphi)$ et de droite Ka.

Supposons maintenant E de dimension $n \geqslant 1$. Complétons $a \in H$ en une base $(e_1, \ldots, e_{n-1})$ de $\mathrm{Ker}(\varphi)$, de façon à ce que $e_{n-1} = a$. Soit $e_n \in E$ tel que $\varphi(e_n) = 1$. Alors, $\mathbf{e} = (e_1, \ldots, e_n)$ est une base de E (tout cela est possible en vertu du lemme III-5.8).

Par construction, on a

$$u(e_i) = e_i, \quad \text{pour } i \in [\![1, n-1]\!], \quad \text{et} \quad u(e_n) = e_n + a = e_{n-1} + e_n.$$

On en déduit alors que

$$\mathrm{Mat}(u\,;\mathbf{e}) = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & 1 \\ & & & 1 \end{pmatrix}.$$

On a donc (3) $\Longrightarrow$ (4). De plus, (4) $\Longrightarrow$ (5) est évident. Il reste à démontrer l'implication (5) $\Longrightarrow$ (1). Soit $\mathbf{e} = (e_1, \ldots, e_n)$ une base de E telle que

$$\mathrm{Mat}(u\,;\mathbf{e}) = T_{i,j}(\lambda).$$

Posons $H = \mathrm{Vect}_K(e_k, k \neq j)$ et $D = Ke_i$. Alors H est un hyperplan, et D est une droite contenue dans H (puisque $i \neq j$). Des calculs élémentaires montrent alors que u est une transvection d'hyperplan H et de droite D.

Cela achève la démonstration. □

5.3. Remarque. Une transvection u est un élément de $\mathrm{GL}(E)$, et son inverse est encore une transvection. Pour le voir, choisissons une forme linéaire non nulle $\varphi : E \longrightarrow K$ et $a \in \mathrm{Ker}(\varphi)$ non nul tels que

$$u(x) = x + \varphi(x)a, \qquad \text{pour tout } x \in E.$$

On vérifie par un simple calcul que la transvection

$$\begin{aligned} E &\longrightarrow E \\ x &\longmapsto x - \varphi(x)a \end{aligned}$$

est l'inverse de u.

De plus, lorsque E est de dimension finie, le point (4) montre qu'une transvection est en fait un élément de $\mathrm{SL}(E)$.

Le point (5) nous invite à poser la définition suivante.

5.4. Définition. Soit $n \geqslant 1$. Une *matrice de transvection* de $\mathrm{M}_n(K)$ est une matrice de la forme

$$T_{i,j}(\lambda) = \begin{array}{c} \\ i \rightarrow \end{array} \begin{array}{c} \begin{matrix} \\ j \\ \downarrow \end{matrix} \\ \begin{pmatrix} 1 & & & \\ & \ddots & \lambda & \\ & & \ddots & \\ & & & 1 \end{pmatrix} \end{array},$$

où $i \neq j$ et $\lambda \in K^\times$.

Remarquons que $T_{i,j}(\lambda) \in \mathrm{SL}_n(K)$ et que

$$T_{i,j}(\lambda)^{-1} = T_{i,j}(-\lambda).$$

Enfin, si $M \in \mathrm{M}_n(K)$, notons $L_1, \ldots, L_n$ ses lignes, et $C_1, \ldots, C_n$ ses colonnes. Alors, l'opération

$$M \longrightarrow T_{i,j}(\lambda)M$$

revient à faire l'opération élémentaire sur les lignes

$$L_i \longleftarrow L_i + \lambda L_j,$$

tandis que l'opération

$$M \longrightarrow MT_{i,j}(\lambda)$$

revient à faire l'opération élémentaire sur les colonnes

$$C_j \longleftarrow C_j + \lambda C_i.$$

On va maintenant définir les dilatations.

5.5. Définition. Soit E un K-espace vectoriel non nul. Soient H un hyperplan de E et D une droite de E tels que $E = H \oplus D$ et soit $\lambda \in K \setminus \{0, 1\}$. La *dilatation* d'hyperplan H, de direction D et de rapport λ est l'unique endomorphisme $u \in \mathscr{L}(E)$ tel que

(1) on a $u(x) = x$, pour tout $x \in H$;

(2) on a $u(x) = \lambda x$, pour tout $x \in D$.

Autrement dit, on a

$$u(x_H + x_D) = x_H + \lambda x_D,$$

pour tout $x_H \in H$ et tout $x_D \in D$.

On dit que $u \in \mathscr{L}(E)$ est une dilatation s'il existe un hyperplan H et une droite D supplémentaires dans E, et un scalaire $\lambda \in K \setminus \{0, 1\}$ tels que u soit la dilatation d'hyperplan H, de direction D et de rapport λ.

Le résultat suivant caractérise les dilatations.

5.6. Proposition. *Soit E un K-espace vectoriel non nul. Pour tout endomorphisme $u \in \mathscr{L}(E)$, les propriétés suivantes sont équivalentes.*

(1) *L'endomorphisme u est une dilatation.*

(2) *L'endomorphisme u est un automorphisme de E, $\mathrm{Im}(u - \mathrm{Id}_E)$ est une droite, et $\mathrm{Im}(u - \mathrm{Id}_E)$ n'est pas contenu dans $\mathrm{Ker}(u - \mathrm{Id}_E)$.*

Dans ce cas, u est une dilatation d'hyperplan $H = \mathrm{Ker}(u - \mathrm{Id}_E)$ et de droite $D = \mathrm{Im}(u - \mathrm{Id}_E)$, et son rapport λ est l'unique scalaire $\lambda \in K$ tel que $u(e) = \lambda e$, où e est un élément arbitraire de $D \setminus H$.

Si E est de dimension finie, ces conditions sont aussi équivalentes à la propriété suivante.

(3) *Il existe une base de E dans laquelle la matrice représentative de u est de la forme*

$$D(\lambda) = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \lambda \end{pmatrix},$$

où $\lambda \in K \setminus \{0, 1\}$.

Démonstration

(1) $\Longrightarrow$ (2). Soit $u \in \mathscr{L}(E)$ la dilatation d'hyperplan H, de droite D et de rapport λ. Il est clair que la dilatation d'hyperplan H, de droite D et de rapport λ^{-1} est l'inverse de u, donc $u \in \mathrm{GL}(E)$. Un calcul explicite montre aussi que $\mathrm{Ker}(u - \mathrm{Id}_E) = H$ et $\mathrm{Im}(u - \mathrm{Id}_E) = D$. Puisque $E = H \oplus D$, D n'est pas contenue dans H, d'où le résultat.

(2) $\Longrightarrow$ (1). Supposons que $u \in \mathscr{L}(E)$ vérifie les conditions indiquées. Posons $D = \mathrm{Im}(u - \mathrm{Id}_E)$ et $H = \mathrm{Ker}(u - \mathrm{Id}_E)$. Puisque D est une droite, on a $D = Ke$, avec $e \in D$ non nul. De plus, on a $e \notin H$, car sinon D serait contenue dans H. Puisque $u(e) - e \in D$, il existe $\alpha \in K$ tel que

$$u(e) - e = \alpha e.$$

Remarquons que $\alpha \neq 0$ puisque $e \notin H$.

Montrons maintenant que $E = H \oplus D$. Soit $x \in E$. Par définition de D et de e, il existe $\mu \in K$ tel que $u(x) - x = \mu e$. Posons $x_H = x - \mu\alpha^{-1}e$. On a

$$u(x_H) = x + \mu e - \mu\alpha^{-1}(1 + \alpha)e = x - \mu\alpha^{-1}e = x_H,$$

et ainsi $x_H \in H$. On a donc $x = x_H + \mu\alpha^{-1}e$, avec $x_H \in H$, d'où $E = H + D$.

Si maintenant $x \in H \cap D$, alors $u(x) - x = 0$ d'une part, et $x = \mu e$, $\mu \in K$. Mais alors, on a $0 = u(x) - x = \mu\big(u(e) - e\big) = \mu\alpha e$. Comme e est non nul et que $\alpha \neq 0$, on en déduit que $\mu = 0$, soit $x = 0$. Finalement, $E = H \oplus D$.

Par définition, $u(x) = x$ pour tout $x \in H$. De plus, pour tout $\mu \in K$, on a

$$u(\mu e) = \mu u(e) = \mu(1 + \alpha)e = (1 + \alpha)(\mu e).$$

Remarquons que $1 + \alpha \neq 1$, puisque $\alpha \neq 0$. De plus, $1 + \alpha \neq 0$, sinon on aurait $u(e) = 0$ et par suite $e = 0$ par injectivité de u, ce qui contredit le choix de e. Finalement, u est la dilatation d'hyperplan H, de droite D et de rapport $1 + \alpha$.

Supposons E de dimension finie n.

En recollant une base de H et une base de D, on obtient une base de E dans laquelle la matrice représentative de e est $D(\lambda)$. Inversement, s'il existe une base $\mathbf{e} = (e_1, \ldots, e_n)$ telle que $\mathrm{Mat}(u\,; \mathbf{e}) = D(\lambda)$ avec $\lambda \in K \setminus \{0, 1\}$, alors

il est clair que u est la dilatation d'hyperplan $H = \mathrm{Vect}_K(e_1, \ldots, e_{n-1})$, de droite $D = Ke_n$ et de rapport λ, d'où l'équivalence (1) $\iff$ (3).

Cela achève la démonstration. □

5.7. Définition. Soit $n \geqslant 1$. Une *matrice de dilatation* de $\mathrm{M}_n(K)$ est une matrice de la forme

$$D(\lambda) = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \lambda \end{pmatrix},$$

où $\lambda \in K \setminus \{0, 1\}$.

Remarquons que $D(\lambda) \in \mathrm{GL}_n(K)$ et que $D(\lambda)^{-1} = D(\lambda^{-1})$.

Enfin, si $M \in \mathrm{M}_n(K)$, l'opération

$$M \longrightarrow D(\lambda)M$$

revient à faire l'opération élémentaire sur les lignes

$$L_n \longleftarrow \lambda L_n,$$

tandis que l'opération

$$M \longrightarrow MD(\lambda)$$

revient à faire l'opération élémentaire sur les colonnes

$$C_n \longleftarrow \lambda C_n.$$

Nous allons maintenant donner quelques résultats classiques sur la structure des groupes $\mathrm{GL}(E)$ et $\mathrm{SL}(E)$.

5.8. Proposition. *Soit K un corps et soit $n \geqslant 1$.*

(1) *Toute matrice $M \in \mathrm{SL}_n(K)$ est produit de matrices de transvection ;*

(2) *toute matrice $M \in \mathrm{GL}_n(K)$ est produit de matrices de transvection et d'au plus une matrice de dilatation.*

Autrement dit, si E est un K-espace vectoriel non nul de dimension finie, on a

(1′) *tout $u \in \mathrm{SL}(E)$ est la composée de transvections ;*

(2′) *tout $u \in \mathrm{GL}(E)$ est la composée de transvections et d'au plus une dilatation.*

Démonstration. Les propriétés (1′) et (2′) sont des conséquences directes de (1) et (2). Il suffit donc de démontrer ces dernières.

Soit $M \in \mathrm{SL}_n(K)$. Nous allons montrer qu'il existe des matrices de transvection $T_1, \ldots, T_r, T'_1, \ldots, T'_s$ telles que

$$T_1 \cdots T_r M T'_1 \cdots T'_s = \mathrm{I}_n,$$

ce qui suffira à démontrer la première assertion, en utilisant le fait que l'inverse d'une matrice de transvection est encore une matrice de transvection. Cela revient à dire que l'on peut passer de M à I_n en faisant des opérations du type

$$C_j \longleftarrow C_j + \lambda C_i\,, \quad L_i \longleftarrow L_i + \lambda L_j, \quad i \neq j.$$

On procède par récurrence sur n.

Si $n = 1$, il n'y a rien à faire. Supposons que le résultat soit vrai pour toute matrice de $\mathrm{SL}_{n-1}(K)$ et soit $M = (a_{ij}) \in \mathrm{SL}_n(K)$. Puisque M est en particulier inversible, L_1 est non nulle, donc il existe j tel que $a_{1j} \neq 0$. Si $j \neq 1$ et si $a_{11} = 0$, on effectue

$$C_1 \longleftarrow C_1 + C_j$$

et l'on peut donc supposer $a_{11} \neq 0$. Si maintenant $a_{21} = 0$, on effectue

$$L_2 \longleftarrow L_2 + L_1$$

et l'on peut supposer que l'on a aussi $a_{21} \neq 0$. On fait ensuite l'opération

$$L_1 \longleftarrow L_1 + \frac{1 - a_{11}}{a_{21}} L_2,$$

pour obtenir une matrice telle que $a_{11} = 1$. En faisant alors des opérations du type

$$C_j \longleftarrow C_j + \lambda C_1\,, \quad L_j \longleftarrow L_j + \lambda L_1,$$

on obtient une matrice diagonale par blocs de la forme $M'' = \begin{pmatrix} 1 & 0 \\ 0 & M' \end{pmatrix}$.

Faire des opérations élémentaires ne change pas le déterminant, donc cette matrice est aussi de déterminant 1. En développant par rapport à la première colonne, on obtient $\det(M') = 1$, c'est-à-dire $M' \in \mathrm{SL}_{n-1}(K)$. Par hypothèse de récurrence, on peut faire des opérations sur les lignes et les colonnes de M' du type précédent pour la ramener à I_{n-1}. Il existe donc des matrices de transvection $P_1, \ldots, P_k, P'_1, \ldots, P'_\ell \in \mathrm{SL}_{n-1}(K)$ telles que

$$P_1 \cdots P_k M' P'_1 \cdots P'_\ell = \mathrm{I}_{n-1}.$$

Posons alors

$$T_i = \begin{pmatrix} 1 & 0 \\ 0 & P_i \end{pmatrix}, \quad T'_j = \begin{pmatrix} 1 & 0 \\ 0 & P'_j \end{pmatrix}.$$

Ce sont des matrices de transvections, et le calcul par blocs montre alors que $T_1 \cdots T_k M'' T'_1 \cdots T'_\ell = \mathrm{I}_n$.

Comme M'' est obtenue en multipliant M à gauche et à droite par des matrices de transvections, on a le résultat escompté. Cela achève alors

la récurrence. Pour montrer le dernier point, il suffit de remarquer que si $M \in \mathrm{GL}_n(K) \setminus \mathrm{SL}_n(K)$, alors $MD\big(\det(M)\big)^{-1} \in \mathrm{SL}_n(K)$, et d'appliquer le premier point. □

On continue ce chapitre en démontrant que les dilatations suffisent pour décrire les éléments de $\mathrm{GL}(E)$.

5.9. Proposition. *Soit K un corps possédant au moins trois éléments, et soit E un K-espace vectoriel non nul de dimension finie. Alors, tout élément de $\mathrm{GL}(E)$ est la composée de dilatations.*

Démonstration. Soit $n = \dim_K(E)$.

Si $n = 1$, tout élément de $\mathrm{GL}(E)$ différent de Id_E est une dilatation (d'hyperplan $\{0\}$ et de droite E). Si $n \geqslant 2$, il suffit de montrer par la proposition précédente que toute transvection est un produit de dilatations.

Soit $u \in \mathrm{SL}(E)$ une transvection. Par la proposition 5.2, en page 76, il existe une base $\mathbf{e} = (e_1, \ldots, e_n)$ telle que

$$\mathrm{Mat}(u\,;\mathbf{e}) = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & 1 \\ & & & 1 \end{pmatrix}.$$

Par hypothèse, il existe[5] $\lambda \in K \setminus \{0, 1\}$. Remarquons maintenant que l'on a l'égalité

$$\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & 1 \\ & & & 1 \end{pmatrix} = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \lambda^{-1} \end{pmatrix} \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & 1 \\ & & & \lambda \end{pmatrix}.$$

La première matrice est la matrice représentative de la dilatation d'hyperplan $H = \mathrm{Vect}_K(e_1, \ldots, e_{n-1})$, de droite $D = Ke_n$ et de rapport λ^{-1} dans la base $\mathbf{e}$. Nous allons maintenant interpréter la deuxième matrice comme la matrice représentative d'une dilatation bien choisie dans cette même base.

Remarquons que l'on a l'identité

$$P \begin{pmatrix} 1 & 0 \\ 0 & \lambda \end{pmatrix} P^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & \lambda \end{pmatrix}, \quad \text{avec } P = \begin{pmatrix} 1 & 1 \\ 0 & \lambda - 1 \end{pmatrix}.$$

5. Tout corps contient au moins les éléments distincts 0 et 1. Un éventuel troisième élément de K est donc nécessairement différent de ceux là.

Ainsi, si l'on pose $Q = \begin{pmatrix} \mathrm{I}_{n-2} & \\ & P \end{pmatrix} \in \mathrm{GL}_n(K)$, on a

$$QD(\lambda)Q^{-1} = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & 1 \\ & & & \lambda \end{pmatrix}.$$

En interprétant Q comme la matrice de passage d'une certaine base $\mathbf{f}$ à $\mathbf{e}$, on a le résultat. Si l'on veut être plus précis, par définition de la matrice Q, on a $\mathbf{f} = (e_1, \ldots, e_{n-1}, e_{n-1} + (\lambda - 1)e_n)$, et la matrice $\begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & 1 \\ & & & \lambda \end{pmatrix}$ est donc la matrice représentative dans la base $\mathbf{e}$ de la dilatation d'hyperplan H, de droite $D' = K\big(e_{n-1} + (\lambda - 1)e_n\big)$ et de rapport λ, comme on peut le vérifier par calcul direct.

Ainsi, u est produit de deux dilatations, et l'on a fini. □

5.10. Remarque. Il faut faire un peu attention à ce que dit cet énoncé, et surtout à ce qu'il ne dit pas. Une fois que l'on a fixé une base de E, les matrices représentatives des dilatations correspondantes n'ont aucune raison des matrices de la forme $D(\lambda)$, $\lambda \in K^\times$.

Un des problèmes majeurs de l'algèbre linéaire en dimension finie est de savoir si sous quelles conditions deux endomorphismes donnés sont semblables ou non. Ce problème sera résolu dans l'avant-dernière partie du livre et nécessite une étude approfondie de la réduction des endomorphismes. Néanmoins, on peut commencer par essayer de résoudre le problème pour une classe spécifique d'endomorphismes. Les transvections et les dilatations jouant un rôle important comme on vient de le voir, on peut légitimement se poser la question suivante : à quelles conditions deux transvections (ou deux dilatations) sont-elles semblables dans $\mathrm{GL}(E)$?

Comme les transvections sont des éléments de $\mathrm{SL}(E)$, on peut aussi raffiner la question et se demander si elles sont semblables dans $\mathrm{SL}(E)$, i.e. si l'on peut trouver un changement de base de déterminant 1 qui permet de passer de l'une à l'autre. On a en fait le résultat suivant.

5.11. Proposition. *Soit E un K-espace vectoriel de dimension $n \geqslant 2$, et soit $\mathbf{e}$ une base de E. Pour $\lambda \in K^\times$, on note $u_\lambda \in \mathrm{SL}(E)$ la transvection dont la matrice représentative dans la base $\mathbf{e}$ est*

$$U_\lambda = \begin{pmatrix} \mathrm{I}_{n-2} & & \\ & 1 & \lambda \\ & 0 & 1 \end{pmatrix}.$$

Alors :

(1) *toute transvection est semblable à u_1 dans $\mathrm{GL}(E)$. En particulier, toutes les transvections sont semblables dans $\mathrm{GL}(E)$;*

(2) *si $n \geqslant 3$, toute transvection est semblable à u_1 dans $\mathrm{SL}(E)$. En particulier, toutes les transvections sont semblables dans $\mathrm{SL}(E)$;*

(3) *si $n = 2$, toute transvection est semblable dans $\mathrm{SL}(E)$ à une transvection u_λ, $\lambda \in K^\times$. De plus, pour tous $\lambda, \mu \in K^\times$, u_λ et u_μ sont semblables dans $\mathrm{SL}(E)$ si, et seulement si, $\lambda\mu^{-1}$ est un carré dans K.*

Démonstration. Pour alléger les notations, si $u, v \in \mathscr{L}(E)$, on notera uv l'application linéaire $u \circ v \in \mathscr{L}(E)$.

Le point (1) est clair par la proposition 5.2 (5), puisque toute transvection admet une matrice représentative égale à U_1 dans une base bien choisie. Montrons le point (2). Soit $u \in \mathrm{SL}(E)$ une transvection. D'après ce qui précède, il existe $v \in \mathrm{GL}(E)$ telle que $u = vu_1v^{-1}$. Soit $w \in \mathrm{GL}(E)$ l'unique endomorphisme de E dont la matrice représentative dans la base $\mathbf{e}$ est

$$W = \begin{pmatrix} \mathrm{I}_{n-3} & & & \\ & \det(v) & & \\ & & \det(v)^{-1} & \det(v)^{-1} \\ & & & \det(v)^{-1} \end{pmatrix}.$$

Cela est possible, car $n \geqslant 3$. On a alors

$$\det(w) = \det(W) = \det(v)^{-1},$$

et l'on a $WU_1W^{-1} = U_1$. Par conséquent, $\det(vw) = 1$, et ainsi $vw \in \mathrm{SL}(E)$. Mais alors, on a

$$(vw)u_1(vw)^{-1} = v(wu_1w^{-1})v^{-1} = vu_1v^{-1} = u.$$

Il reste à démontrer le point (3). Soit $u \in \mathrm{SL}(E)$ une transvection arbitraire, et soit $U = \mathrm{Mat}(u\,;\mathbf{e})$. Par (1), il existe $P \in \mathrm{GL}_2(K)$ tel que

$$PUP^{-1} = U_1 = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Soit $\lambda = \det(P)^{-1}$, et posons $Q = \begin{pmatrix} \lambda & \\ & 1 \end{pmatrix} P$. Alors, $Q \in \mathrm{SL}_2(K)$, et l'on a

$$QUQ^{-1} = \begin{pmatrix} \lambda & \\ & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \lambda^{-1} & \\ & 1 \end{pmatrix} = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} = U_\lambda.$$

L'unique endomorphisme $v \in \mathscr{L}(E)$ tel que $\mathrm{Mat}(v\,;\mathbf{e}) = Q$ est alors un élément de $\mathrm{SL}(E)$ vérifiant $vuv^{-1} = u_\lambda$.

Enfin, soient λ, $\mu \in K^\times$. Si $\lambda = \mu\gamma^2$, $\gamma \in K^\times$, on a

$$\begin{pmatrix} \gamma^{-1} & \\ & \gamma \end{pmatrix} \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \gamma^{-1} & \\ & \gamma \end{pmatrix}^{-1} = \begin{pmatrix} 1 & \mu \\ 0 & 1 \end{pmatrix},$$

ce qui démontre que u_λ et u_μ sont semblables dans $\mathrm{SL}(E)$. Inversement, supposons qu'il existe $P = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}_2(K)$ telle que $PU_\lambda P^{-1} = U_\mu$. Alors, la relation $PU_\lambda = U_\mu P$ donne

$$a = a + \mu c, \quad \lambda a + b = b + \mu d, \quad \lambda c + d = d.$$

On en tire $c = 0$ et $a = \lambda^{-1}\mu d$. Comme $\det(P) = ad - bc = ad = 1$, on en déduit que $1 = \lambda^{-1}\mu d^2$, et donc que $\lambda\mu^{-1}$ est un carré dans K. □

Le cas des dilatations est beaucoup plus facile et laissé au lecteur en exercice. Toutes les considérations précédentes sur les transvections et les dilations sont essentiellement tirées de [37].

6. Exercices

E1. Soit K un corps, et soient $\alpha_1, \ldots, \alpha_n \in K$. La *matrice de Vandermonde* associée à $(\alpha_1, \ldots, \alpha_n)$ est

$$M(\alpha_1, \ldots, \alpha_n) = (\alpha_i^{j-1})_{i,j} \in \mathrm{M}_n(K).$$

a. Montrer que le déterminant de $M(\alpha_1, \ldots, \alpha_n)$ est égal à $\prod_{i<j} (\alpha_j - \alpha_i)$.

b. En déduire que $M(\alpha_1, \ldots, \alpha_n)$ est inversible si, et seulement si, les scalaires $\alpha_1, \ldots, \alpha_n$ sont distincts.

E2. Soit $n \geqslant 1$ et soit $\zeta_n = e^{\frac{2i\pi}{n}}$. On pose

$$M = (\zeta_n^{(k-1)(\ell-1)})_{k,\ell} \quad \text{et} \quad \overline{M} = (\overline{\zeta}_n^{(k-1)(\ell-1)})_{k,\ell}.$$

a. Vérifier que $M \in \mathrm{GL}_n(\mathbb{C})$ et que $M^{-1} = \frac{1}{n}\overline{M}$. En déduire $|\det(M)|$.

b. Calculer M^2, puis M^4. En déduire que $\det(M) = \pm i^{\frac{(n-2)(n-1)}{2}} n^{\frac{n}{2}}$.

c. En remarquant que M est une matrice de Vandermonde, montrer que l'on a $\det(M) = i^{\frac{(n-2)(3n-2)}{2}} n^{\frac{n}{2}}$. On établira aussi l'égalité

$$\prod_{1 \leqslant k \leqslant \frac{n}{2}} \sin\left(\frac{k\pi}{n}\right) = \frac{\sqrt{n}}{2^{\frac{n-1}{2}}}.$$

E3. Soient $n \geqslant 1$, et $a_0, \ldots, a_{n-1} \in \mathbb{C}$. On note $\mathbb{U}_n = \{z \in \mathbb{C} \mid z^n = 1\}$. On définit

$$P = a_{n-1}X^{n-1} + \cdots + a_1 X + a_0 \in \mathbb{C}[X].$$

On note $J = \begin{pmatrix} O_{1\times(n-1)} & \mathrm{I}_{n-1} \\ 1 & O_{(n-1)\times 1} \end{pmatrix} \in \mathrm{M}_n(\mathbb{C})$.

On pose également

$$A = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_{n-1} & a_0 & a_1 & \cdots & a_{n-2} \\ a_{n-2} & a_{n-1} & a_0 & \cdots & a_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & a_3 & \cdots & a_0 \end{pmatrix}, \quad B = \begin{pmatrix} a_0 & a_1 & a_2 & \cdots & a_{n-1} \\ a_1 & a_2 & a_3 & \cdots & a_0 \\ a_2 & a_3 & a_4 & \cdots & a_1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n-1} & a_0 & a_1 & \cdots & a_{n-2} \end{pmatrix}.$$

On dit que A est une matrice *circulante*, et que B est une matrice *anticirculante*.

a. Exprimer A au moyen des puissances de J.

b. Calculer $M^{-1}JM$, où M est définie dans l'exercice précédent.

c. En déduire l'expression de la matrice $M^{-1}AM$ et l'égalité

$$\det(A) = \prod_{z\in\mathbb{U}_n} P(z).$$

d. Soit $p \geqslant 1$, et soient U, V, $W \in \mathrm{M}_p(\mathbb{C})$ commutant deux à deux. Montrer l'égalité

$$\begin{vmatrix} U & V & W \\ W & U & V \\ V & W & U \end{vmatrix} = \det(U^3 + V^3 + W^3 - 3UVW).$$

e. Vérifier que $BM = \left(\zeta_n^{-(k-1)(\ell-1)} P(\zeta_n^{\ell-1})\right)_{k,\ell}$, puis calculer MBM.

f. En déduire que $\det(B) = (-1)^{\frac{(n-1)(n-2)}{2}} \prod_{z\in\mathbb{U}_n} P(z)$.

g. Retrouver ce résultat en calculant M^2B et en utilisant la question **c.**

E4. Soit $n \geqslant 1$. Calculer le déterminant de l'application linéaire

$$\begin{aligned} \mathrm{M}_n(\mathbb{C}) &\longrightarrow \mathrm{M}_n(\mathbb{C}) \\ M &\longmapsto M^t. \end{aligned}$$

E5. Soit Ω un ensemble non vide, soit K un corps et soit $n \geqslant 1$. Finalement, soient $f_1, \ldots, f_n \in \mathscr{F}(\Omega, K)$.

Montrer par récurrence sur n que la famille $(f_1, \dots, f_n)$ est libre si, et seulement si, il existe $x_1, \dots, x_n \in \Omega$ tel que $\det\left(\left(f_i(x_j)\right)_{i,j}\right) \neq 0$.

E6. Soient $n \geqslant 1$, et $a_1, a_2, \dots, a_n \in K^\times$. Soit $A = (a_{\min(i,j)})_{i,j} \in \mathrm{M}_n(K)$.

a. Calculer $\det(A)$ et A^{-1} (dans le cas où A est inversible).

b. Application numérique : $a_k = (-1)^k$, $k \in [\![1, n]\!]$.

E7. Soient $n \geqslant 1$, et $b_1, \dots, b_n \in K$. Soit $B \in \mathrm{M}_n(K)$ la matrice tridiagonale symétrique

$$B = \begin{pmatrix} b_1 + b_2 & -b_2 & & & \\ -b_2 & b_2 + b_3 & -b_3 & & \\ & -b_3 & \ddots & \ddots & \\ & & \ddots & b_{n-1} + b_n & -b_n \\ & & & -b_n & b_n \end{pmatrix}.$$

Calculer $\det(B)$ et B^{-1} (lorsque B est inversible).

E8. Calculer les déterminants suivants.

a. $\det\left((a_i + \delta_{ij} b_j)_{i,j}\right)$, avec $a_1, \dots, a_n, b_1, \dots, b_n \in K$.

b. $$\begin{vmatrix} 1+x^2 & -x & & & \\ -x & 1+x^2 & -x & & \\ & -x & 1+x^2 & -x & \\ & & \ddots & \ddots & -x \\ & & & -x & 1+x^2 \end{vmatrix}.$$

E9. Montrer que deux dilatations de $\mathrm{GL}(E)$ sont semblables si, et seulement si, elles ont même rapport.

E10. Soit F un sous-espace de $\mathrm{M}_n(\mathbb{C})$ contenant au moins une matrice inversible A. On veut montrer que F possède une base constituée de matrices inversibles [6].

On note $\mathscr{I}$ l'ensemble des matrices inversibles contenues dans F.

6. Le lecteur connaissant déjà les propriétés de base des polynômes à coefficients dans un corps quelconque s'apercevra que le résultat reste vrai si l'on remplace $\mathbb{C}$ par un corps possédant au moins $n+1$ éléments.

a. Justifier qu'il suffit de démontrer que $F = \mathrm{Vect}_{\mathbb{C}}(\mathscr{I})$.

b. On suppose qu'il existe $B \in F \setminus \mathrm{Vect}_{\mathbb{C}}(\mathscr{I})$.

(i) Montrer que $\det(XA + B) \in \mathbb{C}(X)$ est un polynôme de degré n à coefficients complexes, de coefficient dominant $\det(A)$.

(ii) En déduire qu'il existe $t_0 \in \mathbb{C}$ tel que $\det(t_0 A + B) \neq 0$.

(iii) En déduire que $B \in \mathrm{Vect}_{\mathbb{C}}(\mathscr{I})$ et une contradiction. Conclure.

Chapitre V

Un peu de géométrie

Comme nous le verrons dans les chapitres ultérieurs, géométrie et théorie des groupes sont intimement liés. C'est pourquoi nous concluons cette partie en introduisant quelques notions de géométrie, ce qui nous permettra au passage de définir quelques groupes matriciels.

1. Affinités vectorielles

Nous allons ici étudier très brièvement les affinités vectorielles, qui englobent des familles d'endomorphismes que le lecteur aura déjà peut-être croisés (projections, symétries, réflexions, dilatations).

1.1. Définition. Soit E un K-espace vectoriel, et soient F et G deux sous-espaces de E tels que $E = F \oplus G$. Enfin, soit $\lambda \in K$. L'*affinité* de base F, de direction G et de rapport λ est l'unique endomorphisme $u \in \mathscr{L}(E)$ tel que

(1) on a $u(x) = x$, pour tout $x \in F$;

(2) on a $u(x) = \lambda x$, pour tout $x \in G$;

Autrement dit, on a

$$u(x_F + x_G) = x_F + \lambda x_G,$$

pour tout $x_F \in F$ et tout $x_G \in G$.

On dit que $u \in \mathscr{L}(E)$ est une *affinité* s'il existe deux sous-espaces supplémentaires F et G et un scalaire $\lambda \in K$ tels que u soit l'affinité de base F, de direction G et de rapport λ.

1.2. Exemples. Le lecteur a déjà rencontré beaucoup d'affinités, sans le savoir peut-être. Soit u l'affinité de base F, de direction G et de rapport λ.

(1) Si $\lambda = 1$, ou si $G = \{0\}$, alors $u = \mathrm{Id}_E$.

(2) Si $\lambda = 0$, alors u est la projection sur F parallèlement à G, c'est-à-dire l'unique endomorphisme $u \in \mathscr{L}(E)$ tel que
   (i) $u(x) = x$ pour tout $x \in F$;
   (ii) $u(x) = 0$ pour tout $x \in G$.

(3) Si $\lambda = -1$, alors u est la symétrie par rapport à F, parallèlement à G, c'est-à-dire l'unique endomorphisme $u \in \mathscr{L}(E)$ tel que
   (i) $u(x) = x$ pour tout $x \in F$;
   (ii) $u(x) = -x$ pour tout $x \in G$.

   En particulier, si F est un hyperplan et G est une droite, u est la réflexion d'hyperplan F et de direction G.

(4) Si $F = \{0\}$ (et donc $G = E$), alors u est l'homothétie de rapport λ.

(5) Si F est un hyperplan et G est une droite, alors u est la dilatation d'hyperplan F, de direction G et de rapport λ (lorsque $\lambda \neq 0, 1$).

Le résultat suivant caractérise les affinités.

1.3. Théorème. *Soit $u \in \mathscr{L}(E)$. Les propriétés suivantes sont équivalentes :*

(1) *l'endomorphisme u est une affinité ;*

(2) *il existe $\lambda \in K \setminus \{1\}$ tel que $(u - \mathrm{Id}_E) \circ (u - \lambda \mathrm{Id}_E) = 0$.*
   Dans ce cas, l'endomorphisme u est l'affinité de base $F = \mathrm{Ker}(u - \mathrm{Id}_E)$, de direction $G = \mathrm{Ker}(u - \lambda \mathrm{Id}_E)$ et de rapport λ.

Si E est de dimension finie, ces deux conditions sont aussi équivalentes à la propriété suivante :

(3) *il existe une base de E dans laquelle la matrice représentative de u est diagonale par blocs, de la forme*

$$\begin{pmatrix} \mathrm{I}_r & \\ & \lambda \mathrm{I}_{n-r} \end{pmatrix},$$

   où $r \in [\![0, n]\!]$ et $\lambda \in K$.

Démonstration

(1) $\Longrightarrow$ (2). Soit $u \in \mathscr{L}(E)$ l'affinité de base F, de direction G et de rapport λ. Pour $x = x_F + x_G \in E$, on a

$$(u - \lambda \mathrm{Id}_E)(x) = (1 - \lambda) x_F.$$

Par conséquent,

$$\big((u-\mathrm{Id}_E)\circ(u-\lambda\mathrm{Id}_E)\big)(x)=(u-\mathrm{Id}_E)\big((1-\lambda)x_F\big)=(1-\lambda)(u-\mathrm{Id}_E)(x_F)=0.$$

Ainsi, $(u-\mathrm{Id}_E)\circ(u-\lambda\mathrm{Id}_E)=0$. Si $\lambda\neq 1$, on a fini. Si $\lambda=1$, on a $u=\mathrm{Id}_E$, d'où $u-\mathrm{Id}_E=0$. En particulier, $(u-\mathrm{Id}_E)\circ u=0$.

(2) $\Longrightarrow$ (1). Nous allons montrer que $E=\mathrm{Ker}(u-\mathrm{Id}_E)\oplus\mathrm{Ker}(u-\lambda\mathrm{Id}_E)$.

Puisque $\lambda\neq 1$, on a l'égalité

$$\mathrm{Id}_E=-\frac{1}{1-\lambda}(u-\mathrm{Id}_E)+\frac{1}{1-\lambda}(u-\lambda\mathrm{Id}_E). \qquad (*)$$

Pour tout $x\in E$, on pose

$$x_\lambda=-\frac{1}{1-\lambda}(u-\mathrm{Id}_E)(x) \quad \text{et} \quad x_1=\frac{1}{1-\lambda}(u-\lambda\mathrm{Id}_E)(x).$$

Remarquons que l'on a

$$(u-\mathrm{Id}_E)\circ(u-\lambda\mathrm{Id}_E)=u\circ u-(1+\lambda)u+\lambda\mathrm{Id}_E=(u-\lambda\mathrm{Id}_E)\circ(u-\mathrm{Id}_E).$$

Ces égalités et l'hypothèse montrent facilement que $x_1\in\mathrm{Ker}(u-\mathrm{Id}_E)$ et que $x_\lambda\in\mathrm{Ker}(u-\lambda\mathrm{Id}_E)$. D'après $(*)$, on a $x=x_1+x_\lambda$. Par ailleurs, si l'on a $x\in\mathrm{Ker}(u-\mathrm{Id}_E)\cap\mathrm{Ker}(u-\lambda\mathrm{Id}_E)$, l'égalité $(*)$ implique alors que l'on a $x=\mathrm{Id}_E(x)=0$, d'où le résultat annoncé.

Posons alors $F=\mathrm{Ker}(u-\mathrm{Id}_E)$ et $G=\mathrm{Ker}(u-\lambda\mathrm{Id}_E)$. On vient de montrer que $E=F\oplus G$, et par définition de F et G, on a $u(x)=x$ si $x\in F$, ainsi que $u(x)=\lambda x$ si $x\in G$, d'où le résultat voulu.

Supposons maintenant E de dimension finie.

(1) $\Longrightarrow$ (3). En recollant une base de F et une base de G, on obtient une base de E dans laquelle la matrice représentative de u à la forme voulue.

(3) $\Longrightarrow$ (1). Soit $\mathbf{e}=(e_1,\ldots,e_n)$ une base de E telle que

$$\mathrm{Mat}(u\,;\mathbf{e})=\begin{pmatrix}\mathrm{I}_r & \\ & \lambda\mathrm{I}_{n-r}\end{pmatrix},$$

avec $r\in[\![0,n]\!]$ et $\lambda\in K$.

Posons $F=\mathrm{Vect}_K(e_1,\ldots,e_r)$ et $G=\mathrm{Vect}_K(e_{r+1},\ldots,e_n)$. Alors, on a l'égalité $E=F\oplus G$ et il est alors clair que u est l'affinité de base F, de direction G et de rapport λ. Cela achève la démonstration. □

Nous allons maintenant appliquer le théorème précédent pour retrouver des résultats bien connus sur les projecteurs et les symétries.

1.4. Définition. On dit que $p \in \mathscr{L}(E)$ est un *projecteur* (ou une projection) s'il existe deux sous-espaces F et G supplémentaires tels que p soit la projection sur F parallèlement à G.

On dit que $s \in \mathscr{L}(E)$ est une *symétrie* s'il existe deux sous-espaces F et G de E supplémentaires tels que p soit la symétrie par rapport à F parallèlement à G.

Le théorème 1.3 donne alors immédiatement le résultat suivant.

1.5. Corollaire. *Soit E un K-espace vectoriel.*

(1) *Un endomorphisme $p \in \mathscr{L}(E)$ est un projecteur si, et seulement si, il vérifie $p \circ p = p$. Dans ce cas, l'endomorphisme p est la projection sur $\mathrm{Ker}(\mathrm{Id}_E - p) = \mathrm{Im}(p)$ parallèlement à $\mathrm{Ker}(p)$.*

Si de plus E est de dimension finie, cela équivaut à l'existence d'une base de E dans laquelle la matrice représentative de u est diagonale par blocs, de la forme

$$\begin{pmatrix} \mathrm{I}_r & \\ & O_{n-r} \end{pmatrix},$$

où $r \in [\![0, n]\!]$.

(2) *Un endomorphisme $s \in \mathscr{L}(E)$ est une symétrie si, et seulement si, il vérifie $s \circ s = \mathrm{Id}_E$. Dans ce cas, l'endomorphisme s est la symétrie par rapport à $\mathrm{Ker}(s - \mathrm{Id}_E)$ parallèlement à $\mathrm{Ker}(s + \mathrm{Id}_E)$.*

Si de plus E est de dimension finie, cela équivaut à l'existence d'une base de E dans laquelle la matrice représentative de u est diagonale par blocs, de la forme

$$\begin{pmatrix} \mathrm{I}_r & \\ & -\mathrm{I}_{n-r} \end{pmatrix},$$

où $r \in [\![0, n]\!]$.

Démonstration. Puisque l'on a les équivalences

$$p \circ p = p \iff (p - \mathrm{Id}_E) \circ p = 0,$$

et

$$s \circ s = \mathrm{Id}_E \iff (s - \mathrm{Id}_E) \circ (s + \mathrm{Id}_E) = 0,$$

tout découle du théorème 1.3, à part l'égalité $\mathrm{Ker}(\mathrm{Id}_E - p) = \mathrm{Im}(p)$, que nous démontrons maintenant. Si $x \in \mathrm{Im}(p)$, il existe un vecteur $x' \in E$ tel que $x = p(x')$. On a alors

$$(\mathrm{Id}_E - p)(x) = (\mathrm{Id}_E - p)\big(p(x')\big) = (p - p \circ p)(x') = 0.$$

Inversement, si $x \in \mathrm{Ker}(\mathrm{Id}_E - p)$, alors on a $x = p(x) \in \mathrm{Im}(p)$. □

2. Espaces euclidiens, espaces hermitiens

2.1. Définition. Soit $K = \mathbb{R}$ ou $\mathbb{C}$, et soit E un K-espace vectoriel. Un *produit scalaire* sur E est une application $\langle\ ,\ \rangle : E \times E \longrightarrow K$ telle que, pour tous $x, x', y, y' \in E$ et tout $\lambda \in K$, on a

(1) $\langle x + x', y\rangle = \langle x, y\rangle + \langle x', y\rangle$ et $\langle x, y + y'\rangle = \langle x, y\rangle + \langle x, y'\rangle$;

(2) $\langle \lambda x, y\rangle = \overline{\lambda}\langle x, y\rangle$ et $\langle x, \lambda y\rangle = \lambda\langle x, y\rangle$;

(3) $\langle y, x\rangle = \overline{\langle x, y\rangle}$.
En particulier, $\langle x, x\rangle \in \mathbb{R}$, pour tout $x \in E$;

(4) pour tout $x \in E$, $\langle x, x\rangle \geqslant 0$;

(5) pour tout $x \in E$, $\langle x, x\rangle = 0 \implies x = 0$.

Remarquons que lorsque $K = \mathbb{R}$, les conditions (1) – (3) se résument simplement en disant que l'application $\langle\ ,\ \rangle$ est une forme bilinéaire symétrique, puisque l'on a $\overline{\lambda} = \lambda$, pour tout $\lambda \in \mathbb{R}$.

Deux vecteurs $x, y \in E$ sont dits *orthogonaux* pour le produit scalaire $\langle\ ,\ \rangle$ si $\langle x, y\rangle = 0$.

Deux sous-espaces F et G d'un espace euclidien E sont dit *orthogonaux* si pour tout $x \in F$ et tout $y \in G$, on a $\langle x, y\rangle = 0$.

Une famille de sous-espaces $(F_i)_{i\in I}$ de E est dite en *somme directe orthogonale* si pour tous $i, j \in I$, $i \neq j$, F_i et F_j sont orthogonaux. Il est facile de voir que ces sous-espaces sont alors en somme directe. On le note

$$\overset{\perp}{\bigoplus_{i\in I}} F_i.$$

Une famille de vecteurs $(x_i)_{i\in I}$ de E est dite *orthonormée* si elle vérifie

$$\langle x_i, x_j\rangle = \delta_{ij}, \quad \text{pour tous } i, j \in I.$$

Si F est un sous-espace vectoriel de E, on définit *l'orthogonal* $F^\perp$ de F par

$$F^\perp = \{x \in E \mid \langle x, y\rangle = 0, \text{ pour tout } y \in F\}.$$

C'est un sous-espace vectoriel de E.

2.2. Définition. Un couple $(E, \langle\ ,\ \rangle)$, où E est un K-espace vectoriel **de dimension finie** et $\langle\ ,\ \rangle$ est un produit scalaire est appelé un espace *euclidien* si $K = \mathbb{R}$ et un espace *hermitien* si $K = \mathbb{C}$.

On dira souvent « soit E un espace euclidien/hermitien » en oubliant de mentionner le produit scalaire.

Notation. Si $M = (a_{ij}) \in \mathrm{M}_{m\times n}(K)$, on note $M^* = (\overline{a}_{ji})$. En particulier, si $K = \mathbb{R}$, on a $M^* = M^t$.

2.3. Exemple. Soit $n \geqslant 1$. L'application

$$\begin{aligned} \mathbb{C}^n \times \mathbb{C}^n &\longrightarrow \mathbb{C} \\ (X, Y) &\longmapsto X^*Y \end{aligned}$$

est un produit scalaire sur $\mathbb{C}^n$. De même, l'application

$$\begin{aligned} \mathbb{R}^n \times \mathbb{R}^n &\longrightarrow \mathbb{C} \\ (X, Y) &\longmapsto X^tY \end{aligned}$$

est un produit scalaire sur $\mathbb{R}^n$.

Dans les deux cas, on l'appellera le produit scalaire usuel (ou standard).

Le résultat suivant rappelle quelques propriétés des espaces euclidiens et hermitiens.

2.4. Proposition. *Soit $(E, \langle\ ,\ \rangle)$ un espace euclidien ou hermitien. Alors :*

(1) *toute famille orthonormée de E est libre, et se complète en une base orthonormée. En particulier, tout sous-espace de E admet une base orthonormée (i.e. une base qui soit une famille orthonormée) ;*

(2) *si $\mathbf{e}$ est une base orthonormée, alors pour tous x, $y \in E$, on a*

$$\langle x, y\rangle = [x]_{\mathbf{e}}^*\,[y]_{\mathbf{e}}$$

(3) *pour tout sous-espace F de E, on a $(F^\perp)^\perp = F$, et $E = F \overset{\perp}{\oplus} F^\perp$. En particulier, $\dim(F^\perp) = \dim(E) - \dim(F)$;*

(4) *pour tout $u \in \mathscr{L}(E)$, il existe un unique $u^* \in \mathscr{L}(E)$ tel que*

$$\langle u(x), y\rangle = \langle x, u^*(y)\rangle, \qquad \text{pour tout } y \in E.$$

Si $\mathbf{e}$ est une base **orthonormée** *de E, alors on a*

$$\mathrm{Mat}(u^*\,;\mathbf{e}) = \mathrm{Mat}(u\,;\mathbf{e})^*.$$

En particulier, $\det(u^*) = \overline{\det(u)}$.

De plus, $\mathrm{Id}_E^* = \mathrm{Id}_E$, et pour tous u_1, $u_2 \in \mathscr{L}(E)$, on a

$$(u_1 \circ u_2)^* = u_2^* \circ u_1^*.$$

Tous ces résultats seront redémontrés dans un cadre un plus général dans le chapitre XXXII.

2.5. Remarque. Si F est un espace euclidien ou hermitien, il possède un supplémentaire naturel, à savoir $F^\perp$, et l'on peut alors définir la projection *orthogonale* sur F (resp. la symétrie *orthogonale* par rapport à F) comme étant la projection sur F (resp. la symétrie par rapport à F) parallèlement au sous-espace $F^\perp$.

2.6. Définition. Si E est un espace euclidien ou hermitien, et si $u \in \mathscr{L}(E)$, l'endomorphisme $u^* \in \mathscr{L}(E)$ s'appelle l'endomorphisme *adjoint* à u, ou l'adjoint de u.

On dit que $u \in \mathscr{L}(E)$ est

(1) un endomorphisme *normal* si $u^* \circ u = u \circ u^*$;

(2) une *isométrie* si l'on a [1] $u^* \circ u = \mathrm{Id}_E$;

(3) un endomorphisme *autoadjoint* si $u^* = u$;

(4) un endomorphisme *antiautoadjoint* si $u^* = -u$.

On remarque que les isométries, ainsi que les endomorphismes adjoints ou antiautoadjoints sont normaux.

Lorsque $K = \mathbb{R}$, on parle aussi d'endomorphisme symétrique ou antisymétrique, plutôt que d'endomorphisme autoadjoint ou antiautoadjoint.

L'ensemble des isométries a une structure particulièrement agréable.

2.7. Définition. L'ensemble des isométries de E forme un sous-groupe [2] de $\mathrm{GL}(E)$, que l'on note parfois $\mathrm{Iso}(E)$. C'est exactement l'ensemble des endomorphismes de E qui conservent le produit scalaire.

Lorsque $K = \mathbb{R}$, le groupe $\mathrm{Iso}(E)$ est plutôt noté $\mathrm{O}(E)$, et appelé le *groupe orthogonal* de E .

Lorsque $K = \mathbb{C}$, le groupe $\mathrm{Iso}(E)$ est plutôt noté $\mathrm{U}(E)$, et appelé le *groupe unitaire* de E . Dans ce cas, une isométrie est aussi appelé un endomorphisme *unitaire.*

Nous allons maintenant introduire une définition supplémentaire, afin de pouvoir retraduire ce qui précède en termes matriciels.

---

1. Dans ce cas, puisque l'on est en dimension finie, u est un automorphisme de E, et l'on a $u^{-1} = u^*$.

2. Cela se démontre facilement en utilisant le dernier point de la proposition précédente.

2.8. Définition. On dit qu'une matrice $M \in \mathrm{M}_n(\mathbb{C})$ est *hermitienne* si l'on a $M^* = M$ et *antihermitienne* si l'on a $M^* = -M$.

Une matrice $M \in \mathrm{M}_n(\mathbb{R})$ sera dite *symétrique* si l'on a $M^t = M$ et *antisymétrique* si l'on a $M^t = -M$.

Une matrice $M \in \mathrm{M}_n(\mathbb{C})$ est dite *unitaire* si $M^*M = \mathrm{I}_n$.

L'ensemble $\mathrm{U}_n(\mathbb{C})$ des matrices unitaires est un sous-groupe de $\mathrm{GL}_n(\mathbb{C})$.

Une matrice $M \in \mathrm{M}_n(\mathbb{R})$ est dite *orthogonale* si l'on a $M^tM = \mathrm{I}_n$.

L'ensemble $\mathrm{O}_n(\mathbb{R})$ des matrices orthogonales est un sous-groupe de $\mathrm{GL}_n(\mathbb{R})$.

2.9. Remarques

(1) Une matrice $P \in \mathrm{M}_n(K)$ est unitaire/orthogonale si, et seulement si, ses colonnes forment une base orthonormée de K^n pour le produit scalaire usuel.

En effet, P est unitaire/orthogonale si, et seulement si, ses colonnes forment une famille orthonormée de cardinal n par définition de P^* et du produit matriciel, ou de manière équivalente une base orthonormée (puisque une famille orthonormée est toujours libre).

(2) Soit $u \in \mathscr{L}(E)$ d'un endomorphisme d'un espace hermitien.

Si $M \in \mathrm{M}_n(\mathbb{C})$ est la matrice représentative de u dans une base **orthonormée**, alors u est :

(i) normal si, et seulement si, $M^*M = MM^*$;
(ii) une isométrie si, et seulement si, M est unitaire ;
(iii) autoadjoint si, et seulement si, M est hermitienne ;
(iv) antiautoadjoint si, et seulement si, M est antihermitienne.

(3) Soit $u \in \mathscr{L}(E)$ d'un endomorphisme d'un espace euclidien.

Si $M \in \mathrm{M}_n(\mathbb{R})$ est la matrice représentative de u dans une base **orthonormée**, alors u est :

(i) normal si, et seulement si, $M^tM = MM^t$;
(ii) une isométrie si, et seulement si, M est orthogonale ;
(iii) symétrique si, et seulement si, M est symétrique ;
(iv) antisymétrique si, et seulement si, M est antisymétrique.

Remarquons au passage que les matrices unitaires et orthogonale nous permettent de caractériser les bases orthonormées.

2.10. Lemme. *Soit* $\mathbf{e} = (e_1, \ldots, e_n)$ *une base orthonormée d'un espace euclidien (resp. hermitien)* E *et soit* $\mathbf{e}' = (e'_1, \ldots, e'_n)$ *une base de* E. *Alors,* $\mathbf{e}'$ *est une base orthonormée de* E *si, et seulement si, la matrice de passage de* $\mathbf{e}'$ *à* $\mathbf{e}$ *est orthogonale (resp. unitaire).*

Démonstration. Soit P la matrice de passage de $\mathbf{e}'$ à $\mathbf{e}$. Alors, P sera orthogonale (resp. unitaire) si, et seulement si, la famille $([e'_1]_{\mathbf{e}}, \ldots, [e'_n]_{\mathbf{e}})$ est une base orthonormée de K^n pour le produit scalaire usuel, Cela d'après la remarque 2.9 (1). Puisque $\mathbf{e}$ est orthonormée, cela revient exactement à dire que

$$\langle e'_i, e'_j \rangle = \delta_{ij}, \quad \text{pour tous } i, j \in [\![1, n]\!],$$

c'est-à-dire que $\mathbf{e}'$ est une famille orthonormée de cardinal n, i.e. une base orthonormée. □

Dans ce chapitre, nous étudierons surtout les isométries d'un espace euclidien ou hermitien.

Donnons maintenant quelques exemples d'isométries.

2.11. Exemples

(1) Dans le plan euclidien, les rotations vectorielles de $\mathbb{R}^2$, ainsi que les symétries orthogonales par rapport à une droite passant par l'origine, sont des isométries.

(2) Si E est un espace euclidien ou hermitien, alors une symétrie s de E est une isométrie si, et seulement si, c'est une symétrie orthogonale par rapport à un sous-espace F de E. Dans ce cas, s est la symétrie orthogonale par rapport à $F = \mathrm{Ker}(s - \mathrm{Id}_E)$, et l'on a $F^\perp = \mathrm{Ker}(s + \mathrm{Id}_E)$.

En effet, supposons tout d'abord que s soit la symétrie orthogonale par rapport à F. En recollant une base orthonormée de F et une base orthonormée de $F^\perp$, on obtient une base orthonormée $\mathbf{e}$ de E telle que

$$M = \mathrm{Mat}(s\,; \mathbf{e}) = \begin{pmatrix} \mathrm{I}_r & \\ & -\mathrm{I}_{n-r} \end{pmatrix},$$

où n et r sont les dimensions de E et F. Ainsi, $M^*M = \mathrm{I}_n$, et $s \in \mathrm{U}(E)$.

Inversement, supposons que la symétrie s soit une isométrie. D'après le corollaire 1.5, s est la symétrie par rapport à $F = \mathrm{Ker}(s - \mathrm{Id}_E)$ parallèlement à $G = \mathrm{Ker}(s + \mathrm{Id}_E)$.

Nous allons vérifier que $G = F^\perp$, ce qui donnera la conclusion. Tout d'abord, on a

$$\dim(G) = \dim(E) - \dim(F) = \dim(F^\perp),$$

donc il suffit de vérifier que $G \subset F^\perp$.

Soit $x \in G$. Pour tout $y \in F$, on a

$$\langle s(x), s(y) \rangle = \langle -x, y \rangle = -\langle x, y \rangle.$$

Mais, s étant une isométrie, on a aussi

$$\langle s(x), s(y) \rangle = \langle x, y \rangle.$$

On en déduit que $\langle x, y\rangle = 0$. Cela étant vrai pour tout $y \in F$, on a bien $x \in F^\perp$, d'où la conclusion.

Le lemme suivant caractérise les isométries en termes d'images des éléments d'une base orthonormée.

2.12. Lemme. *Soit E un espace euclidien ou hermitien de dimension n, et soit $u \in \mathscr{L}(E)$. Les propriétés suivantes sont équivalentes :*

(1) *l'endomorphisme u est une isométrie ;*

(2) *pour toute base orthonormée $(e_1, \dots, e_n)$ de E, $\big(u(e_1), \dots, u(e_n)\big)$ est une base orthonormée de E ;*

(3) *il existe au moins une base orthonormée $(e_1, \dots, e_n)$ de E telle que la famille $\big(u(e_1), \dots, u(e_n)\big)$ soit une base orthonormée de E.*

Démonstration

(1) $\Longrightarrow$ (2). Supposons que u soit une isométrie et soit $(e_1, \dots, e_n)$ une base orthonormée de E. Pour tous i, $j \in [\![1, n]\!]$, on a alors

$$\langle u(e_i), u(e_j)\rangle = \langle e_i, e_j\rangle = \delta_{ij}.$$

La famille $\big(u(e_1), \dots, u(e_n)\big)$ est donc orthonormée. En particulier, elle est libre, et comme elle est de cardinal n, c'est une base de E.

(2) $\Longrightarrow$ (3). Évident.

(3) $\Longrightarrow$ (1). Soit $\mathbf{e} = (e_1, \dots, e_n)$ une base orthonormée de E telle que la famille $\mathbf{e}' = \big(u(e_1), \dots, u(e_n)\big)$ soit une base orthonormée de E.

Alors, la matrice de passage de $\mathbf{e}'$ à $\mathbf{e}$ est unitaire/orthogonale d'après le lemme 2.10. Mais, cette matrice de passage n'est autre que $\mathrm{Mat}(u\,;\mathbf{e})$. Par la remarque 2.9 (2), u est une isométrie. □

Pour finir ce paragraphe, nous allons établir un résultat qui montre que les endomorphismes normaux se comporte bien vis-à-vis des sous-espaces stables.

2.13. Lemme. *Soit E un espace euclidien ou hermitien, soit $u \in \mathscr{L}(E)$ et soit F un sous-espace vectoriel de E stable par u. Si u est respectivement un endomorphisme normal, autoadjoint, antiautoadjoint ou une isométrie, alors F et $F^\perp$ sont stables par u et u^*, et les endomorphismes u_F, $u_{F^\perp}$, u_F^* et $u_{F^\perp}^*$ sont respectivement normaux, autoadjoints, antiautoadjoints ou des isométries.*

Démonstration. Bien sûr, on peut supposer F non nul. Choisissons une base orthonormée de F, et une base orthonormée de $F^\perp$. En recollant ces

deux bases, on obtient une base orthonormée $\mathbf{e}$ de E. Soit M la matrice représentative de u dans cette base. Comme F est stable par u, on a

$$M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix},$$

et la matrice représentative de u^* est

$$M^* = \begin{pmatrix} A^* & 0 \\ B^* & C^* \end{pmatrix}.$$

Si u vérifie une des propriétés de l'énoncé, on a dans tous les cas

$$M^*M = MM^*,$$

ce qui se traduit par l'égalité

$$\begin{pmatrix} A^*A & A^*B \\ B^*A & B^*B + C^*C \end{pmatrix} = \begin{pmatrix} AA^* + BB^* & BC^* \\ CB^* & CC^* \end{pmatrix}.$$

En particulier, $A^*A = AA^* + BB^*$. En appliquant la trace, et en utilisant le fait que $\mathrm{tr}(MN) = \mathrm{tr}(NM)$ pour toutes matrices carrées M, N, on obtient l'égalité $\mathrm{tr}(BB^*) = 0$. Or, si $B = (b_{ij})$, on a $\mathrm{tr}(BB^*) = \sum_{i,j} |b_{ij}|^2$. Ainsi, on en déduit que $b_{ij} = 0$, pour tous i, j, c'est-à-dire $B = 0$.

Par construction de la base $\mathbf{e}$, on en déduit que F et $F^\perp$ sont stables par u et u^*. Il s'ensuit que A, C, A^* et C^* sont les matrices de u_F, $u_{F^\perp}$, u_F^* et $u_{F^\perp}^*$ dans les bases orthonormées de F et $F^\perp$ choisies au départ. Si maintenant on a $MM^* = M^*M$, $M^* = M, M^* = -M$ ou $M^*M = \mathrm{I}_n$, il en est de même de A, A^*, C et C^*, d'où le résultat. □

3. Isométries d'un espace euclidien : premiers résultats

Dans ce paragraphe, E est un espace euclidien.

Remarquons que si $u \in \mathrm{O}(E)$, alors on a $1 = \det(u^*)\det(u) = \det(u)^2$, et donc $\det(u) = \pm 1$.

3.1. Définition. Une isométrie $u \in \mathrm{O}(E)$ d'un espace euclidien E est dite *directe* si l'on a $\det(u) = 1$, et *indirecte* si l'on a $\det(u) = -1$. L'ensemble des isométries directes forment un sous-groupe de $\mathrm{O}(E)$, noté $\mathrm{O}^+(E)$ ou encore $\mathrm{SO}(E)$, appelé *groupe spécial orthogonal*. L'ensemble des isométries indirectes est noté parfois $\mathrm{O}^-(E)$. Ce n'est pas un sous-groupe de $\mathrm{O}(E)$.

Nous allons maintenant étudier d'un peu plus près la structure de $\mathrm{O}^\pm(E)$ lorsque $\dim_{\mathbb{R}}(E) = 2$ ou 3. Pour comprendre ces ensembles, il suffit de

décrire les matrices représentatives de leurs éléments dans une base orthonormée.

Notation. Si $n \geqslant 1$, on pose

$$\mathrm{O}_n^+(\mathbb{R}) = \{M \in \mathrm{O}_n(\mathbb{R}) \mid \det(M) = 1\}$$

et

$$\mathrm{O}_n^-(\mathbb{R}) = \{M \in \mathrm{O}_n(\mathbb{R}) \mid \det(M) = -1\}.$$

Si E est un espace euclidien de dimension n, une isométrie u de E est directe (resp. indirecte) si, et seulement si, sa matrice représentative dans une base orthonormée est dans $\mathrm{O}_n^+(\mathbb{R})$ (resp. $\mathrm{O}_n^-(\mathbb{R})$).

Un calcul direct montre le résultat suivant.

3.2. Proposition. *On a les égalités suivantes :*

$$(1)\quad \mathrm{O}_2^+(\mathbb{R}) = \left\{\begin{pmatrix} a & -b \\ b & a \end{pmatrix} \mid a^2+b^2=1,\ a,b\in\mathbb{R}\right\}$$
$$= \left\{\begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix} \mid \theta \in \mathbb{R}\right\}.$$

$$(2)\quad \mathrm{O}_2^-(\mathbb{R}) = \left\{\begin{pmatrix} a & b \\ b & -a \end{pmatrix} \mid a^2+b^2=1, a,b\in\mathbb{R}\right\}$$
$$= \left\{\begin{pmatrix} \cos(\theta) & \sin(\theta) \\ \sin(\theta) & -\cos(\theta) \end{pmatrix} \mid \theta \in \mathbb{R}\right\}.$$

L'interprétation géométrique est la suivante : si E un plan euclidien, les isométries directes de E sont les rotations planes, et les isométries indirectes sont les symétries orthogonales par rapport à une droite.

3.3. Remarque. Si E est un espace euclidien de dimension 2, et si u est une isométrie indirecte, il existe une base orthonormée de E pour laquelle la matrice de u dans cette base est

$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Cela se comprend bien géométriquement : les vecteurs parallèles à l'axe de symétrie sont fixes, et les vecteurs orthogonaux à l'axe sont envoyés sur leurs opposés. Il suffit donc de prendre un vecteur unitaire de l'axe de symétrie, et un vecteur unitaire orthogonal à cet axe.

On passe au cas de la dimension 3.

3.4. Proposition. *Soit E un espace euclidien de dimension 3, et soit u une isométrie de E.*

(1) *Si $u \in \mathrm{O}^+(E)$, alors il existe une base orthonormée de E dans laquelle la matrice représentative de u est de la forme*

$$\begin{pmatrix} 1 & & \\ & \cos(\theta) & -\sin(\theta) \\ & \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

(2) *Si $u \in \mathrm{O}^-(E)$, alors il existe une base orthonormée de E dans laquelle la matrice représentative de u est de la forme*

$$\begin{pmatrix} -1 & & \\ & \cos(\theta) & -\sin(\theta) \\ & \sin(\theta) & \cos(\theta) \end{pmatrix}.$$

Cela sera la conséquence directe d'un résultat plus général qui sera démontré plus loin. L'interprétation géométrique est la suivante : si E est un espace euclidien de dimension 3, les éléments de $\mathrm{O}^+(E)$ sont les rotations de E, et ceux de $\mathrm{O}^-(E)$ sont obtenues en effectuant une rotation dans un plan de $\mathbb{R}^3$, puis une réflexion orthogonale par rapport à ce plan.

Nous verrons dans les exercices comme trouver l'axe et l'angle de la rotation.

On va maintenant montrer qu'il existe des isométries indirectes en toute dimension.

3.5. Définition. Soit E un espace euclidien, et soit $x \in E \setminus \{0\}$.

La *réflexion orthogonale* d'hyperplan $H = (\mathbb{R}x)^\perp$ est la symétrie orthogonale par rapport à H (et de direction $H^\perp = \mathbb{R}x$), c'est-à-dire l'unique endomorphisme $\tau_x \in \mathscr{L}(E)$ vérifiant :

(1) $(\tau_x)_{|H} = \mathrm{Id}_H$;

(2) $\tau_x(x) = -x$.

Autrement dit, pour tout $\lambda \in K$, et tout $y \in H$, on a

$$\tau_x(\lambda x + y) = -\lambda x + y.$$

3.6. Lemme. *Pour tout $x \in E \setminus \{0\}$, on a $\tau_x \in \mathrm{O}^-(E)$. De plus, pour tout vecteur $v \in E$, on a*

$$\tau_x(v) = v - 2\frac{\langle x, v\rangle}{\langle x, x\rangle}x.$$

Enfin, $\tau_x^2 = \mathrm{Id}_E$, et $\dim_{\mathbb{R}}(\mathrm{Ker}((\tau_x + \mathrm{Id}_E)) = 1$.

Démonstration. D'après l'exemple 2.11 (2), $\tau_x \in \mathrm{O}(E)$, et il existe une base orthonormée de E dans laquelle la matrice représentative de τ_x est

$$M = \begin{pmatrix} -1 & 0 \\ 0 & \mathrm{I}_{n-1} \end{pmatrix}.$$

Ainsi, τ_x est une isométrie indirecte, puisque $\det(M) = -1$. On en déduit également que l'on a

$$\dim_{\mathbb{R}} \big(\mathrm{Ker}(\tau_x + \mathrm{Id}_E)\big) = \dim_{\mathbb{R}}(M + \mathrm{I}_n) = 1.$$

Pour montrer le second point, il suffit de vérifier que l'application

$$\begin{aligned} E &\longrightarrow E \\ v &\longmapsto v - 2\frac{\langle x, v\rangle}{\langle x, x\rangle} x \end{aligned}$$

est linéaire, envoie x sur $-x$ et se restreint à l'identité sur $(\mathbb{R}x)^\perp$, ce qui est clair. Le fait que $\tau_x^2 = \mathrm{Id}_E$ est évident, puisque τ_x est une symétrie. □

On peut maintenant énoncer le théorème suivant.

3.7. Théorème. *Soit E un espace euclidien de dimension n, et soit u une isométrie de E.*
Alors :

(1) *u est la composée de $n - \dim_{\mathbb{R}} \big(\mathrm{Ker}(u - \mathrm{Id}_E)\big)$ réflexions orthogonales ;*

(2) *si u est la composée de p réflexions orthogonales, alors*

$$p \geqslant n - \dim_{\mathbb{R}} \big(\mathrm{Ker}(u - \mathrm{Id}_E)\big).$$

En particulier, toute isométrie est le produit d'au plus n réflexions orthogonales.

Démonstration. Soit $u \in \mathrm{O}(E)$, et soit $F_u = \mathrm{Ker}(u - \mathrm{Id}_E)$. Supposons que l'on ait $u = \tau_{x_1} \circ \cdots \circ \tau_{x_p}$, et posons $F = \mathrm{Vect}_{\mathbb{R}}(x_1, \ldots, x_p)$. Par définition, on a $\dim_{\mathbb{R}}(F) \leqslant p$. Si $x \in F^\perp$, alors x est orthogonal à x_i, et ainsi $\tau_{x_i}(x) = x$ pour tout $i \in [\![1, p]\!]$. Il s'ensuit que $u(x) = x$ pour tout $x \in F^\perp$, d'où l'inclusion $F^\perp \subset F_u$. On a ainsi $\dim_{\mathbb{R}}(F_u) \geqslant \dim_{\mathbb{R}}(F^\perp)$. Or,

$$\dim_{\mathbb{R}}(F^\perp) = n - \dim_{\mathbb{R}}(F) \geqslant n - p,$$

d'où $p \geqslant n - \dim_{\mathbb{R}}(F_u)$.

Démontrons maintenant que u est la composée de $n - \dim_{\mathbb{R}}(F_u)$ réflexions orthogonales. D'après ce qui précède, il suffit de démontrer que u peut s'écrire comme la composée d'au plus $n - \dim_{\mathbb{R}}(F_u)$ réflexions orthogonales.

Nous allons procéder par récurrence sur $p_u = n - \dim_{\mathbb{R}}(F_u)$.

Plus précisément, soit (H_k) la proposition suivante.

(H_k) Toute isométrie $u \in \mathrm{O}(E)$ telle que $p_u \in [\![0, k]\!]$ est la composée d'au plus p_u réflexions orthogonales.

Si $k = 0$, alors $p_u = 0$ et donc $u = \mathrm{Id}_E$. Dans ce cas, u est le produit de 0 réflexions orthogonales. Ainsi, (H_0) est vraie.

Soit $k \in [\![0, n-1]\!]$. Supposons que (H_k) soit vraie, et montrons que (H_{k+1}) est vraie. Il reste à montrer que le résultat attendu est vrai lorsque $u \in \mathrm{O}(E)$ vérifie $p_u = k + 1 \geqslant 1$.

L'idée est trouver une réflexion orthogonale τ telle que $p_{\tau \circ u} < p_u$. Dans ce cas, $\tau \circ u$ sera un produit d'au plus $p_{\tau \circ u}$ réflexions orthogonales par hypothèse de récurrence. Mais alors, l'isométrie $u = \tau \circ (\tau \circ u)$ sera produit d'au plus $p_{\tau \circ u} + 1 \leqslant p_u$ réflexions, ce qui achèvera la récurrence. En fait, on va trouver une réflexion orthogonale τ telle que $F_u \subsetneq F_{\tau \circ u}$. On aura alors $\dim_{\mathbb{R}}(F_u) < \dim_{\mathbb{R}}(F_{\tau \circ u})$ et donc $p_{\tau \circ u} < p_u$. On doit donc exhiber une réflexion orthogonale vérifiant cette propriété. Puisque le sous-espace $F_u^{\perp}$ est de dimension $p_u \geqslant 1$, il existe $x \in F_u^{\perp} \setminus \{0\}$. On a donc $x \neq u(x)$, puisque l'on a $F_u \cap F_u^{\perp} = \{0\}$. Posons $y = u(x)$. Comme F_u est un sous-espace vectoriel stable par u, il en est de même de son orthogonal $F_u^{\perp}$, d'après le lemme 2.13. Ainsi, on a $y \in F_u^{\perp}$, et de plus $y \neq x$, d'où $x - y \in F_u^{\perp} \setminus \{0\}$. Remarquons enfin que l'on a

$$\langle x - y, x + y \rangle = \langle x, x \rangle - \langle y, y \rangle = \langle x, x \rangle - \langle u(x), u(x) \rangle = 0,$$

car u est une isométrie. Cela illustre simplement le fait que les diagonales d'un losange sont orthogonales. On a alors

$$\tau_{x-y}(y) = -\frac{1}{2}\tau_{x-y}(x-y) + \frac{1}{2}\tau_{x-y}(x+y) = -\frac{1}{2}(y-x) + \frac{1}{2}(x+y) = x.$$

Géométriquement, cela s'explique par le fait qu'une diagonale d'un losange est la bissectrice de l'angle formé par deux côtés adjacents.

On va montrer que $\tau = \tau_{x-y}$ vérifie la propriété souhaitée. Soit $z \in F_u$ (on a donc $u(z) = z$). Comme $x - y \in F_u^{\perp}$, z est orthogonal à $x - y$, et l'on obtient alors

$$\tau_{x-y}\big(u(z)\big) = \tau_{x-y}(z) = z.$$

Ainsi, $F_u \subset F_{\tau_{x-y} \circ u}$, et l'inclusion est stricte, car $x \notin F_u$ mais

$$\tau_{x-y}\big(u(x)\big) = \tau_{x-y}(y) = x,$$

et donc $x \in F_{\tau_{x-y} \circ u}$. Cela achève la récurrence, ainsi que la démonstration. □

On s'intéresse maintenant au groupe spécial orthogonal.

3.8. Définition. Soit E un espace euclidien. On dit que $u \in \mathrm{O}(E)$ est un *renversement* si $u^2 = \mathrm{Id}_E$ et $\mathrm{Ker}(u + \mathrm{Id}_E)$ est de dimension 2.

Par l'exemple 2.11 (2), on a $E = \mathrm{Ker}(u - \mathrm{Id}_E) \overset{\perp}{\oplus} \mathrm{Ker}(u + \mathrm{Id}_E)$, et la matrice représentative de u dans une base orthonormée bien choisie est

$$\begin{pmatrix} \mathrm{I}_{n-2} & 0 \\ 0 & -\mathrm{I}_2 \end{pmatrix},$$

puisque $\mathrm{Ker}(u + \mathrm{Id}_E)$ est de dimension 2. En particulier, un renversement est de déterminant 1, et est donc un élément de $\mathrm{O}^+(E)$. En fait, toute isométrie directe est un produit de renversements.

3.9. Théorème. *Soit E un espace euclidien de dimension n. Si $n \geqslant 3$, toute isométrie directe est le produit d'au plus n renversements.*

Démonstration. Soit $u \in \mathrm{O}^+(E)$ une isométrie directe.

Puisque $\det(u) = 1$ et que le déterminant d'une réflexion est -1, le résultat précédent montre que $u = \tau_1 \circ \cdots \circ \tau_{2p}, 2p \leqslant n$, où chaque τ_i est une réflexion. On peut toujours supposer que $u \neq \mathrm{Id}_E$ (car Id_E est le produit de 0 renversements).

Supposons tout d'abord que $n = 3$. Alors, $u = \tau_1 \circ \tau_2$ est le produit d'exactement 2 réflexions (car $u \neq \mathrm{Id}_E$). Puisque τ_i est une réflexion, sa matrice dans une base orthonormée bien choisie est

$$\begin{pmatrix} -1 & & \\ & 1 & \\ & & 1 \end{pmatrix}$$

comme on l'a déjà remarqué. Il devient alors clair que $-\tau_i$ est un renversement, et l'on a fini, puisque $u = (-\tau_1) \circ (-\tau_2)$.

Supposons maintenant que $n > 3$. D'après les considérations précédentes, il suffit de montrer que le produit de deux réflexions s'écrit comme produit de deux renversements. Soient τ_1, τ_2 deux réflexions orthogonales d'hyperplans respectifs H_1, H_2, et posons $u = \tau_1 \circ \tau_2$. Alors, on a $\dim_{\mathbb{R}}(H_1 \cap H_2) = n-2$ ou $n-1$. En particulier, $H_1 \cap H_2$ contient dans tous les cas un sous-espace de dimension $n-3$, disons F. Puisque F est contenu à la fois dans H_1 et dans H_2, on a $u_F = \mathrm{Id}_F$. Par suite, $F^\perp$ est stable par u par le lemme 2.13. En recollant une base orthonormée de F et une base orthonormée de $F^\perp$, on obtient une base orthonormée de E dans laquelle la matrice représentative de u est

$$M = \begin{pmatrix} \mathrm{I}_{n-3} & \\ & M' \end{pmatrix},$$

où M' est la matrice de $u_{F^\perp}$ dans la base orthonormée de $F^\perp$ choisie.

On en déduit alors que $\det(u_{F^\perp}) = 1$, et la relation $M^tM = \mathrm{I}_n$ implique facilement que $M'^tM' = \mathrm{I}_3$. Ainsi, $u_{F^\perp} \in \mathrm{O}^+(F^\perp)$. Or, $\dim_{\mathbb{R}}(F^\perp) = 3$, donc par le cas précédent, il existe des renversements σ'_1, σ'_2 de $F^\perp$ tels que

$$u_{F^\perp} = \sigma'_1 \circ \sigma'_2.$$

On définit alors $\sigma_i \in \mathscr{L}(E)$ par

$$\sigma_i(x+y) = x + \sigma'_i(y), \qquad \text{pour tout } x \in F,\ y \in F^\perp.$$

On vérifie facilement [3] que σ_i est un renversement et que $u = \sigma_1 \circ \sigma_2$. Cela achève la démonstration. □

3.10. Remarque. Le théorème précédent est faux si $n = 2$. En effet, dans ce cas, le seul renversement de E est $-\mathrm{Id}_E$. Mais alors, n'importe quelle rotation plane d'angle $\theta \in \mathbb{R} \setminus \pi\mathbb{Z}$ est une isométrie directe qui ne peut s'écrire comme produit de renversements.

Pour finir ce paragraphe, nous allons déterminer à quelles conditions deux symétries orthogonales s_1, $s_2 \in \mathrm{O}(E)$ sont orthogonalement semblables, c'est-à-dire $s_2 = u \circ s_1 \circ u^{-1}$, avec $u \in \mathrm{O}(E)$. On a le résultat suivant.

3.11. Théorème. *Soit E un espace euclidien, et soient s_1, $s_2 \in \mathrm{O}(E)$ deux symétries orthogonales.*

Alors, il existe $u \in \mathrm{O}^+(E)$ tel que $s_2 = u \circ s_1 \circ u^{-1}$ si, et seulement si, les sous-espaces $\mathrm{Ker}(s_1 + \mathrm{Id}_E)$ et $\mathrm{Ker}(s_2 + \mathrm{Id}_E)$ ont même dimension.

En particulier, c'est le cas si s_1, s_2 sont deux réflexions orthogonales (resp. deux renversements).

Démonstration. Soient s_1, $s_2 \in \mathrm{O}(E)$. Si $u \in \mathrm{O}^+(E)$ vérifie $s_2 = u \circ s_1 \circ u^{-1}$, on a

$$\mathrm{Ker}(s_2 + \mathrm{Id}_E) = \mathrm{Ker}\big(u \circ (s_1 + \mathrm{Id}_E) \circ u^{-1}\big) = u\big(\mathrm{Ker}(s_1 + \mathrm{Id}_E)\big),$$

et comme u est inversible, les sous-espaces $\mathrm{Ker}(s_1 + \mathrm{Id}_E)$ et $\mathrm{Ker}(s_2 + \mathrm{Id}_E)$ ont même dimension.

Inversement, supposons que $E_1 = \mathrm{Ker}(s_1 + \mathrm{Id}_E)$ et $E_2 = \mathrm{Ker}(s_2 + \mathrm{Id}_E)$ aient même dimension $r \geqslant 0$. Soit $(e_1, \ldots, e_r)$ une base orthonormée de E_1, soit $(f_1, \ldots, f_r)$ une base orthonormée de E_2. Soit $(e_{r+1}, \ldots, e_n)$ une base orthonormée de $E_1^\perp$ et soit $(f_{r+1}, \ldots, f_n)$ une base orthonormée de $E_2^\perp$.

Les familles

$$\mathbf{e} = (e_1, \ldots, e_n) \quad \text{et} \quad \mathbf{f} = (f_1, \ldots, f_n)$$

sont alors deux bases orthonormées de E. Soit $u \in \mathscr{L}(E)$ l'unique endo-

3. Cela est laissé en exercice au lecteur.

morphisme de E tel que

$$u(e_i) = f_i, \quad \text{pour tout } i \in [\![1, n]\!].$$

Puisque u envoie une base orthonormée sur une base orthonormée, u est une isométrie. Quitte à remplacer un vecteur de $\mathbf{e}$ par son opposé (ce qui multiplie le déterminant de u par -1, comme on peut le voir matriciellement), on peut supposer que $u \in \mathrm{O}^+(E)$.

Pour tout $i \in [\![1, r]\!]$, on a alors

$$(u \circ s_1 \circ u^{-1})(f_i) = (u \circ s_1 \circ u^{-1})\big(u(e_i)\big) = u\big(s_1(e_i)\big) = u(-e_i) = -f_i = s_2(f_i).$$

D'autre part, puisque s_1 et s_2 sont des symétries orthogonales, on a

$$E_1^{\perp} = \mathrm{Ker}(s_1 - \mathrm{Id}_E) \quad \text{et} \quad E_2^{\perp} = \mathrm{Ker}(s_2 - \mathrm{Id}_E).$$

Par conséquent, pour tout $i \in [\![r+1, n]\!]$, on a

$$(u \circ s_1 \circ u^{-1})(f_i) = (u \circ s_1 \circ u^{-1})\big(u(e_i)\big) = u\big(s_1(e_i)\big) = u(e_i) = f_i = s_2(f_i).$$

Ainsi, $u \circ s_1 \circ u^{-1}$ et s_2 coïncident sur une base de E, et donc $u \circ s_1 \circ u^{-1} = s_2$.

Enfin, une réflexion orthogonale (resp. un renversement) est une symétrie orthogonale s telle que $\mathrm{Ker}(s + \mathrm{Id}_E)$ est de dimension 1 (resp. de dimension 2). Le dernier point est alors clair. Cela achève la démonstration. □

Nous allons maintenant revenir à l'étude des endomorphismes normaux, ce qui nous permettra entre autres d'élucider la structure des isométries d'un espace euclidien ou hermitien quelconque.

4. Structure des endomorphismes normaux

On commence par quelques résultats intermédiaires.

4.1. Lemme. *Soit E un espace euclidien ou hermitien non nul, et soit u un endomorphisme de E.*

(1) *Si E est hermitien, il existe une droite de E stable par u ;*

(2) *si E est euclidien, il existe une droite ou un plan de E stable par u.*

Démonstration. Soit $K = \mathbb{R}$ ou $\mathbb{C}$, et soit $n \geqslant 1$ la dimension de E sur K. Soit M la matrice représentative de u dans une base $\mathbf{e}$ fixée.

Soit $\lambda \in K$. Remarquons tout d'abord que $\mathrm{Ker}(M - \lambda \mathrm{I}_n)$ n'est pas réduit à 0 si, et seulement si, $\lambda \mathrm{I}_n - M$ n'est pas inversible, c'est-à-dire si, et seulement si, $\det(\lambda \mathrm{I}_n - M) = 0$. En utilisant la formule explicite du déterminant, on voit que la quantité $P(\lambda) = \det(\lambda \mathrm{I}_n - M)$ est un polynôme non nul en λ à coefficients dans K.

Le théorème de d'Alembert assure l'existence de $\lambda \in \mathbb{C}$ tel que $P(\lambda) = 0$. Mais alors, d'après ce qui précède, il existe un vecteur $X \in \mathbb{C}^n$ non nul tel que $MX = \lambda X$. Si $K = \mathbb{C}$, la droite engendrée par l'unique vecteur $x \in E$ tel que $[x]_{\mathbf{e}} = X$ est alors stable[4] par u.

Supposons maintenant que $K = \mathbb{R}$, et conservons les notations précédentes. On peut écrire $X = X_0 + iX_1$, avec $X_0, X_1 \in \mathbb{R}^n$ et $\lambda = a + ib$, avec $a, b \in \mathbb{R}$. L'égalité $MX = \lambda X$ se récrit alors

$$MX_0 + iMX_1 = (aX_0 - X_1) + i(bX_0 + aX_1).$$

Comme M est à coefficients réels, par identification, on obtient

$$MX_0 = aX_0 - bX_1 \quad \text{et} \quad MX_1 = bX_0 + aX_1.$$

Soient $x_0, x_1 \in E$ tels que $[x_j]_{\mathbf{e}} = X_j, j = 0, 1$. Alors, le sous-espace de E engendré par x_0 et x_1 est non nul (car X étant non nul, au moins un des deux vecteurs X_0, X_1 est non nul), de dimension $\leqslant 2$, et stable par u. □

4.2. Lemme. *Soit E un plan euclidien*[5]*, et soit $u \in \mathscr{L}(E)$ un endomorphisme normal. Alors, il existe une base orthonormée $\mathbf{e}$ de E dans laquelle la matrice représentative de u dans cette base est, soit diagonale, soit de la forme*

$$\begin{pmatrix} a & -c \\ c & a \end{pmatrix}, \quad a, c \in \mathbb{R},\ c \neq 0.$$

Démonstration

Soit $\mathbf{e}$ une base orthonormée de E, et soit

$$M = \mathrm{Mat}(u\,;\mathbf{e}) = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Puisque u est normal, on a $M^tM = MM^t$, ce qui est équivalent à

$$a^2 + b^2 = a^2 + c^2, \quad ac + bd = ab + cd, \quad b^2 + d^2 = c^2 + d^2.$$

On a donc $c = \varepsilon b$, avec $\varepsilon = \pm 1$. On a alors

$$b(a\varepsilon + d) = b(a + \varepsilon d).$$

Soit $P(\lambda) = \det(\lambda \mathrm{I}_2 - M) = \lambda^2 - (a + d)\lambda + ad - bc$.

Supposons d'abord que $\varepsilon = 1$, i.e. $b = c$. Alors, le discriminant de ce polynôme en λ est

$$(a + d)^2 - 4ad + 4b^2 = (a - d)^2 + 4b^2 \geqslant 0.$$

4. Le lecteur qui connaît sur le bout des doigts la réduction des endomorphismes aura constaté que P n'est rien d'autre que le polynôme caractéristique de u, que λ est une valeur propre de u dans $\mathbb{C}$, et que $x \in E$ est un vecteur propre associé. Nous reviendrons sur ces considérations en détail dans la cinquième partie de cet ouvrage.

5. C'est-à-dire un espace euclidien de dimension 2.

Il existe donc $\lambda \in \mathbb{R}$ tel que $P(\lambda) = 0$, et un raisonnement identique à celui utilisé dans la démonstration précédente montre que E possède au moins une droite F stable par u. Soit e'_1 un vecteur directeur de cette droite, que l'on peut supposer unitaire. Par le lemme 2.13, $F^\perp$ est aussi stable par u. Comme $F^\perp$ est de dimension 1, il est engendré par un vecteur non nul e'_2, que l'on peut supposer unitaire. Alors, $\mathbf{e}' = (e'_1, e'_2)$ est une base orthonormée de E, et la matrice $\mathrm{Mat}(u\,;\mathbf{e}')$ est diagonale (car F et $F^\perp$ sont stables par u et de dimension 1).

Supposons maintenant que $\varepsilon = -1$. Si $b = 0$, alors $c = 0$ et M est diagonale. Si $b \neq 0$, on a $b = -c$ et $-a + d = a - d$, soit $d = a$, ce qui achève la démonstration. □

On alors le résultat suivant.

4.3. Théorème. *Soit $n \geqslant 1$ un entier.*

(1) *Soit E un espace hermitien de dimension n, et soit $u \in \mathscr{L}(E)$ un endomorphisme normal. Alors, il existe une base orthonormée de E dans laquelle la matrice de u est diagonale*[6].

(2) *Soit E un espace euclidien de dimension n, et soit $u \in \mathscr{L}(E)$ un endomorphisme normal. Alors, il existe une base orthonormée de E dans laquelle la matrice de u est de la forme*

$$\begin{pmatrix} \lambda_1 & & & & & & & \\ & \ddots & & & & & & \\ & & \lambda_r & & & & & \\ & & & a_1 & -b_1 & & & \\ & & & b_1 & a_1 & & & \\ & & & & & \ddots & & \\ & & & & & & a_s & -b_s \\ & & & & & & b_s & a_s \end{pmatrix},$$

où r, $s \geqslant 0$ sont des entiers vérifiant $r + 2s = n$, λ_i, $a_j \in \mathbb{R}$, $b_j \in \mathbb{R}^\times$, pour tout $i \in [\![1, r]\!]$, $j \in [\![1, s]\!]$.

Démonstration. On le montre par récurrence sur n. Commençons par le cas hermitien.

Si $n = 1$, il n'y a rien à faire. Supposons que le résultat soit vrai pour tout endomorphisme normal d'un espace hermitien de dimension $n - 1$ et soit u un endomorphisme normal d'un espace hermitien E de dimension n.

6. Dans le langage de la réduction des endomorphismes, dont on parlera ultérieurement, on dira que u est diagonalisable dans une base orthonormée.

Par le lemme 4.1, il existe au moins une droite F de E stable par u. Par le lemme 2.13, $F^\perp$ est stable par u et $u_{F^\perp}$ est normal. Par hypothèse de récurrence, puisque $\dim(F^\perp) = n-1$, il existe une base orthonormée de $F^\perp$ dans laquelle la matrice représentative de $u_{|_{F^\perp}}$ est diagonale. Pour conclure, il suffit prendre un vecteur unitaire dirigeant F et de lui accoler cette base de $F^\perp$ pour obtenir une base orthonormée de E convenable.

Supposons maintenant que E soit un espace euclidien. Encore une fois, il n'y a rien à faire si $n = 1$, et le cas $n = 2$ est donné par le lemme précédent. Supposons que le résultat soit vrai pour tout endomorphisme normal d'un espace euclidien de dimension $\leqslant n-1$, pour un certain $n \geqslant 3$, et soit u un endomorphisme normal d'un espace euclidien E de dimension n.

Par le lemme 4.1, il existe un sous-espace F de E de dimension $\leqslant 2$ stable par u. Par le lemme 2.13, F et $F^\perp$ sont stables par u et les endomorphismes u_F sont $u_{F^\perp}$ sont normaux. Par les cas $n = 1, 2$, il existe une base orthonormée de F dans laquelle la matrice représentative de u_F a la forme voulue. Puisque

$$\dim(F^\perp) = n - \dim(F) \leqslant n-1,$$

par hypothèse de récurrence, il existe une base orthonormée de $F^\perp$ dans laquelle la matrice représentative de $u_{F^\perp}$ a la forme voulue. Pour conclure, il suffit de recoller ces deux bases et de changer l'ordre des vecteurs si nécessaire pour obtenir une base orthonormée de E ayant la propriété requise. □

On peut maintenant élucider la structure des endomorphismes autoadjoints, antiautoadjoints et unitaires d'un espace hermitien.

4.4. Corollaire. *Soit E un espace hermitien.*

(1) *Si $u \in \mathscr{L}(E)$ est autoadjoint, il existe une base orthonormée de E dans laquelle la matrice de u est diagonale à coefficients réels ;*

(2) *si $u \in \mathscr{L}(E)$ est antiautoadjoint, il existe une base orthonormée de E dans laquelle la matrice de u est diagonale à coefficients purement imaginaires ;*

(3) *si $u \in \mathrm{U}(E)$ est unitaire, il existe une base orthonormée de E dans laquelle la matrice de u est diagonale, dont les coefficients diagonaux sont de module 1.*

En particulier,

(i) *pour toute matrice $M \in \mathrm{M}_n(\mathbb{C})$ hermitienne, il existe $P \in \mathrm{U}_n(\mathbb{C})$ telle que P^*MP soit diagonale à coefficients réels ;*

(ii) *pour toute matrice $M \in \mathrm{M}_n(\mathbb{C})$ antihermitienne, il existe $P \in \mathrm{U}_n(\mathbb{C})$ telle que P^*MP soit diagonale à coefficients purement imaginaires ;*

(iii) *pour toute matrice $M \in \mathrm{U}_n(\mathbb{C})$, il existe $P \in \mathrm{U}_n(\mathbb{C})$ telle que P^*MP soit diagonale, dont les coefficients diagonaux sont de module 1.*

Démonstration. Un endomorphisme autoadjoint, antiautoadjoint ou unitaire est normal. On choisit alors une base orthonormée $\mathbf{e}$ comme dans l'énoncé du théorème précédent. La matrice $M = \mathrm{Mat}(u\,;\mathbf{e})$ est alors diagonale. Si u est autoadjoint, antiautoadjoint ou unitaire, M est hermitienne, antihermitienne ou unitaire. Le résultat provient alors d'un simple calcul. Le dernier point est juste une réinterprétation matricielle de ce qui précède. □

Passons maintenant au cas des espaces euclidiens.

4.5. Corollaire. *Soit E un espace euclidien et soit $u \in \mathscr{L}(E)$ un endomorphisme symétrique. Alors, il existe une base orthonormée de E dans laquelle la matrice de u est diagonale à coefficients réels.*

En particulier, pour toute matrice $M \in \mathrm{M}_n(\mathbb{R})$ telle que $M^t = M$, il existe une matrice $P \in \mathrm{O}_n(\mathbb{R})$ telle que P^tMP soit diagonale.

Démonstration. Rappelons qu'un endomorphisme symétrique est normal. On choisit alors une base orthonormée $\mathbf{e}$ comme dans l'énoncé du théorème précédent. Soit $M = \mathrm{Mat}(u\,;\mathbf{e})$. La matrice M est donc symétrique. Mais alors, avec les notations du théorème, on a $s = 0$ dans ce cas (car les b_i sont non nuls) et M est donc diagonale.

Le dernier point est juste une réinterprétation matricielle du résultat. □

4.6. Corollaire. *Soit E un espace euclidien, et soit $u \in \mathscr{L}(E)$ un endomorphisme antisymétrique. Alors, il existe une base orthonormée de E dans laquelle la matrice de u est de la forme*

$$\begin{pmatrix} 0 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & & & \\ & & & 0 & -b_1 & & \\ & & & b_1 & 0 & & \\ & & & & & \ddots & \\ & & & & & 0 & -b_s \\ & & & & & b_s & 0 \end{pmatrix},$$

où $s \geqslant 0$ est un entier et $b_j \in \mathbb{R}^\times$, pour $j \in [\![1, s]\!]$.

Démonstration. Puisque $u^* = -u$, la matrice représentative de u dans n'importe quelle base orthonormée est antisymétrique, en particulier dans la base orthonormée fournie par le théorème précédent. Mais une matrice réelle antisymétrique a des coefficients diagonaux nuls. □

4.7. Corollaire. *Soit E un espace euclidien, et soit $u \in \mathrm{O}(E)$. Alors, il existe une base orthonormée de E dans laquelle la matrice de u est de la forme*

$$\begin{pmatrix} \mathrm{I}_p & & & & & \\ & -\mathrm{I}_q & & & & \\ & & \cos(\theta_1) & -\sin(\theta_1) & & \\ & & \sin(\theta_1) & \cos(\theta_1) & & \\ & & & & \ddots & \\ & & & & \cos(\theta_s) & -\sin(\theta_s) \\ & & & & \sin(\theta_s) & \cos(\theta_s) \end{pmatrix},$$

où p, q, $s \geqslant 0$ vérifient $p+q+2s = \dim_{\mathbb{R}}(E)$, et $\theta_j \in \mathbb{R} \setminus \pi\mathbb{Z}$.

De plus, on a

$$p = \dim_{\mathbb{R}}\big(\mathrm{Ker}(u - \mathrm{Id}_E)\big), \quad q = \dim_{\mathbb{R}}\big(\mathrm{Ker}(u + \mathrm{Id}_E)\big).$$

En particulier, p, q et s sont uniquement déterminés par u.

Démonstration. Puisque $u^* \circ u = \mathrm{Id}_E$, sa matrice représentative dans n'importe quelle base orthonormée est dans $\mathrm{O}_n(\mathbb{R})$, en particulier dans la base fournie par le théorème précédent. Chaque bloc doit être alors dans $\mathrm{O}_1(\mathbb{R})$ ou $\mathrm{O}_2(\mathbb{R})$. Ainsi, on a $\lambda_i = \pm 1$ pour tout $i \in [\![1, r]\!]$ et l'on a $a_j^2 + b_j^2 = 1$ pour tout $j \in [\![1, s]\!]$. On a donc le résultat voulu en utilisant la paramétrisation du cercle-unité, et en réordonnant les vecteurs de base, ce qui donne toujours une base orthonormée. Une rotation plane différente de $\pm\mathrm{Id}_E$ ne stabilisant aucune droite, un simple calcul montre que l'on a

$$p = \dim_{\mathbb{R}}\big(\mathrm{Ker}(u - \mathrm{Id}_E)\big) \quad \text{et} \quad q = \dim_{\mathbb{R}}\big(\mathrm{Ker}(u + \mathrm{Id}_E)\big). \qquad \square$$

4.8. Remarque. Remarquons que l'on a $\det(u) = (-1)^q$. En particulier, si $u \in \mathrm{O}^-(E)$, $q \geqslant 1$. Si de plus $\dim_{\mathbb{R}}(E) = 2$, vu que l'on a $p+q+2s = 2$, on a $s = 0$ et donc $p = 1$, et l'on retrouve le résultat de la remarque 3.3.

Si $\dim_{\mathbb{R}}(E)$ est impair et $u \in \mathrm{O}^+(E)$, alors q est pair, et donc p est impair, car $p+q+2s = n$. En particulier, $p \geqslant 1$. Si de plus E est de dimension 3, on a $(p,q,s) = (1,0,1)$ ou $(p,q,s) = (1,2,0)$. On retrouve alors la première partie de la proposition 3.4, à la formulation près.

Si $u \in \mathrm{O}^-(E)$, alors q est impair, et $(p,q,s) = (2,1,0)$ ou $(0,1,1)$. On retrouve ainsi la seconde partie de la proposition 3.4, encore à la formulation près.

Cela s'interprète comme suit : une isométrie d'un espace euclidien se décompose en produit de réflexions orthogonales et de rotations planes.

5. Espaces euclidiens orientés

Dans ce paragraphe, E désignera un espace vectoriel réel non nul de dimension finie, que l'on ne suppose pas euclidien pour l'instant.

Soient $\mathbf{e} = (e_1, \ldots, e_n)$ et $\mathbf{e}' = (e'_1, \ldots, e'_n)$ deux bases de E. Dans ce qui suit, le réel $\det_{\mathbf{e}}(e'_1, \ldots, e'_n)$ sera noté plus simplement $\det_{\mathbf{e}}(\mathbf{e}')$.

Puisque $\mathbf{e}'$ est une base, $\det_{\mathbf{e}}(\mathbf{e}')$ est non nul. C'est donc un réel qui est, soit strictement positif, soit strictement négatif.

5.1. Définition. On dit que deux bases $\mathbf{e}$ et $\mathbf{e}'$ de E ont *même orientation* si $\det_{\mathbf{e}}(\mathbf{e}') > 0$, et ont des *orientations opposées* sinon.

5.2. Lemme. *Soit E un espace vectoriel réel non nul de dimension finie. La relation « avoir même orientation » est une relation d'équivalence sur l'ensemble des bases de E. De plus, il y a exactement deux classes d'équivalence.*

Démonstration. Soient $\mathbf{e}$, $\mathbf{e}'$, $\mathbf{e}''$ des bases de E. Comme $\det_{\mathbf{e}}(\mathbf{e}) = 1 > 0$, la relation est réflexive. Si maintenant $\det_{\mathbf{e}}(\mathbf{e}') > 0$, alors

$$\det_{\mathbf{e}'}(\mathbf{e}) = \big(\det_{\mathbf{e}}(\mathbf{e}')\big)^{-1} > 0,$$

d'où la symétrie. Enfin, si $\det_{\mathbf{e}}(\mathbf{e}') > 0$ et $\det_{\mathbf{e}'}(\mathbf{e}'') > 0$, alors on a

$$\det_{\mathbf{e}'}(\mathbf{e}'') = \det_{\mathbf{e}'}(\mathbf{e}) \det_{\mathbf{e}}(\mathbf{e}'')$$

par le lemme 2.6, et donc

$$\det_{\mathbf{e}}(\mathbf{e}'') = \det_{\mathbf{e}'}(\mathbf{e}'') \det_{\mathbf{e}}(\mathbf{e}') > 0,$$

d'où la transitivité de la relation.

Montrons maintenant qu'il y a exactement deux classes d'équivalence. Choisissons $\mathbf{e}_0$ une base de E, et soit $\mathbf{e}_1$ une base de E qui ait une orientation opposée à celle de $\mathbf{e}_0$, par exemple la base obtenue en remplaçant le premier vecteur de $\mathbf{e}_0$ par son opposé[7].

Notons que les classes d'équivalence de $\mathbf{e}_0$ et de $\mathbf{e}_1$ sont distinctes, puisque ces deux bases ont des orientations opposées. Il reste alors à montrer que ces deux classes d'équivalence sont les seules, c'est-à-dire que tout base $\mathbf{e}$ de E est orientée dans le même sens qu'une de ces deux bases.

Si $\det_{\mathbf{e}_0}(\mathbf{e}) > 0$, alors $\mathbf{e}_0$ et de $\mathbf{e}$ ont même orientation, et c'est fini. Sinon, on a $\det_{\mathbf{e}_0}(\mathbf{e}) < 0$, et dans ce cas, on a alors

$$\det_{\mathbf{e}_1}(\mathbf{e}) = \det_{\mathbf{e}_1}(\mathbf{e}_0) \det_{\mathbf{e}_0}(\mathbf{e}) > 0,$$

donc $\mathbf{e}_1$ et $\mathbf{e}$ ont même orientation, ce qui achève la démonstration. □

7. Ce qui multiplie le déterminant par -1.

Le lemme précédent nous invite à poser la définition qui suit.

5.3. Définition. Soit E un espace vectoriel réel non nul de dimension finie. Une *orientation* de E est une classe d'équivalence pour la relation précédente.

Il y a donc deux orientations possibles. Néanmoins, en général, aucune des deux orientations n'a de représentant privilégié. Orienter un espace vectoriel réel E revient donc à privilégier une de ces deux classes d'équivalence, ce qui revient à choisir une base de E (et donc l'orientation qui correspond à la classe de la base en question).

5.4. Définition. Un *espace vectoriel orienté* (de dimension finie) est un couple $(E, \mathbf{e}_0)$, où E est un espace vectoriel réel non nul de dimension finie, et $\mathbf{e}_0$ est une base de E.

On dira qu'une base $\mathbf{e}$ d'un espace orienté $(E, \mathbf{e}_0)$ est *directe* si elle est orientée dans le même sens que $\mathbf{e}_0$, c'est-à-dire si $\det_{\mathbf{e}_0}(\mathbf{e}) > 0$, et *indirecte* si elle a une orientation opposée à celle de $\mathbf{e}_0$, c'est-à-dire si $\det_{\mathbf{e}_0}(\mathbf{e}) < 0$.

5.5. Remarque. L'espace $\mathbb{R}^n$ est en général orienté par sa base canonique.

Nous allons maintenant nous concentrer sur le cas des espaces euclidiens. Dans ce cas, nous verrons qu'il est avantageux d'orienter l'espace grâce à une base orthonormée.

On commence par un lemme qui permet de caractériser les bases orthonormées directes.

5.6. Lemme. *Soit E un espace euclidien orienté de dimension $n \geqslant 1$, et soit $\mathbf{e}$ une base orthonormée directe. Pour toute base orthonormée $\mathbf{e}'$, les propositions suivantes sont équivalentes :*

(1) *la base $\mathbf{e}'$ est orthonormée directe ;*

(2) *la matrice de passage de $\mathbf{e}'$ à $\mathbf{e}$ est dans $\mathrm{O}_n^+(\mathbb{R})$;*

(3) $\det_{\mathbf{e}}(\mathbf{e}') = 1$.

Démonstration

(1) $\Longrightarrow$ (2). Par le lemme 2.10, la matrice de passage P de $\mathbf{e}'$ à $\mathbf{e}$ est dans $\mathrm{O}_n(\mathbb{R})$. Comme $P \in \mathrm{O}_n(\mathbb{R})$, on a $\det(P) = \pm 1$. Or, la base $\mathbf{e}'$ étant directe, on a $\det(P) = \det_{\mathbf{e}}(\mathbf{e}') > 0$. Ainsi, $\det(P) = 1$, d'où $P \in \mathrm{O}_n(\mathbb{R})^+$.

(2) $\Longrightarrow$ (3). Par définition même de P, on a $\det_{\mathbf{e}}(\mathbf{e}') = \det(P) = 1$, d'où le résultat.

(3) $\Longrightarrow$ (1). C'est évident, par définition d'une base directe. □

On peut également caractériser les isométries directes d'un espace euclidien orienté.

5.7. Lemme. *Soit E un espace euclidien orienté de dimension $n \geqslant 1$. Pour tout $u \in \mathscr{L}(E)$, les propositions suivantes sont équivalentes :*

(1) *l'endomorphisme u est une isométrie directe ;*

(2) *pour toute base $(e_1, \ldots, e_n)$ orthonormée directe, $\big(u(e_1), \ldots, u(e_n)\big)$ est une base orthonormée directe ;*

(3) *il existe au moins une base orthonormée directe $(e_1, \ldots, e_n)$ telle que la famille $\big(u(e_1), \ldots, u(e_n)\big)$ soit une base orthonormée directe.*

Démonstration

(1) $\Longrightarrow$ (2). Soit $(e_1, \ldots, e_n)$ une base orthonormée directe.

Par le lemme 2.12, $\mathbf{e}' = \big(u(e_1), \ldots, u(e_n)\big)$ est une base orthonormée. Par hypothèse, $\det(u) = 1$, ce qui se traduit avec nos notations par

$$\det_{\mathbf{e}}(\mathbf{e}') = 1.$$

Ainsi, $\mathbf{e}$ et $\mathbf{e}'$ ont même orientation, et comme $\mathbf{e}$ est directe, $\mathbf{e}'$ aussi.

(2) $\Longrightarrow$ (3). Évident.

(3) $\Longrightarrow$ (1). Soit $(e_1, \ldots, e_n)$ une base orthonormée directe telle que la famille $\mathbf{e}' = \big(u(e_1), \ldots, u(e_n)\big)$ soit une base orthonormée directe. Alors, la matrice de passage de $\mathbf{e}'$ à $\mathbf{e}$, qui n'est rien d'autre que $\mathrm{Mat}(u\,;\mathbf{e})$, est un élément de $\mathrm{O}_n^+(\mathbb{R})$ par le lemme précédent, ce qui revient exactement à dire que $u \in \mathrm{O}^+(E)$, d'où le résultat. □

Nous allons maintenant définir la notion de produit mixte. On commence par un lemme.

5.8. Lemme. *Soit E un espace euclidien orienté de dimension $n \geqslant 1$.*

Pour toute base orthonormée directe $\mathbf{e}$, l'application $\det_{\mathbf{e}}$ est indépendante du choix de $\mathbf{e}$.

Démonstration. Soient $\mathbf{e}$ et $\mathbf{e}'$ deux bases orthonormées directes.

On a $\det_{\mathbf{e}'} = \det_{\mathbf{e}'}(\mathbf{e}) \det_{\mathbf{e}}$ par le lemme IV-2.6, et $\det_{\mathbf{e}'}(\mathbf{e}) = 1$ par le lemme 5.6, d'où le résultat. □

5.9. Définition. Soit E un espace euclidien orienté de dimension $n \geqslant 1$. On appelle *produit mixte* l'application $\det_{\mathbf{e}}$, où $\mathbf{e}$ est une base orthonormée directe arbitraire.

Si $x_1, \ldots, x_n \in E$, on note $[x_1, \ldots, x_n]$ le produit mixte de la famille de vecteurs $(x_1, \ldots, x_n)$.

Le résultat qui suit résume les propriétés du produit mixte.

5.10. Lemme. *Soit E un espace euclidien orienté de dimension $n \geqslant 1$. Alors, le produit mixte est une forme n-linéaire alternée sur E. De plus, pour tous $x_1, \ldots, x_n \in E$:*

(1) *on a $[x_1, \ldots, x_n] = 0$ si, et seulement si, la famille $(x_1, \ldots, x_n)$ est liée ;*

(2) *on a $[x_1, \ldots, x_n] > 0$ si, et seulement si, la famille $(x_1, \ldots, x_n)$ est une base directe ;*

Démonstration. Tout découle des propriétés du déterminant et des définitions et résultats précédents. □

5.11. Remarque. Dans tous les résultats démontrés jusqu'à présent, l'espace euclidien E n'est pas nécessairement orienté par une base orthonormée. Néanmoins, orienter un espace euclidien à l'aide d'une base orthonormée $\mathbf{e}_0$ est plutôt un choix naturel : le lemme 5.6 peut alors être appliqué avec $\mathbf{e} = \mathbf{e}_0$, et puisque $\mathbf{e}_0$ est directe, le produit mixte est alors l'application $\det_{\mathbf{e}_0}$.

Cela est particulièrement utile dans le cas où E possède une base orthonormée « évidente », comme par exemple dans le cas de $\mathbb{R}^n$, muni de son produit scalaire naturel, où le choix de la base canonique pour orienter l'espace s'impose de lui-même.

Le lemme suivant est classique.

5.12. Lemme. *Soit $(E, \langle\ ,\ \rangle)$ un espace euclidien de dimension $n \geqslant 1$. Pour toute forme linéaire $\varphi \in E^*$, il existe un unique vecteur $v \in E$ tel que*

$$\varphi(x) = \langle v, x \rangle, \qquad \text{pour tout } x \in E.$$

De plus, si $\mathbf{e} = (e_1, \ldots, e_n)$ est une base orthonormée de E, on a

$$v = \sum_{i=1}^{n} \varphi(e_i) e_i.$$

Démonstration. Soit $\mathbf{e} = (e_1, \ldots, e_n)$ une base orthonormée de E. Posons

$$v = \sum_{i=1}^{n} \varphi(e_i) e_i.$$

Soit $x \in E$, et écrivons $x = \sum_{i=1}^{n} x_i e_i$. Puisque $\mathbf{e}$ est orthonormée, on a alors

$$\langle v, x \rangle = \sum_{i=1}^{n} \varphi(e_i) x_i = \varphi\Big(\sum_{i=1}^{n} x_i e_i \Big) = \varphi(x).$$

Ainsi, l'existence d'un vecteur $v \in E$ vérifiant les conditions du lemme est établie.

Montrons maintenant l'unicité. Supposons qu'il existe v_1, $v_2 \in E$ tels que

$$\varphi(x) = \langle v_1, x\rangle = \langle v_2, x\rangle, \qquad \text{pour tout } x \in E.$$

Alors, pour tout $x \in E$, on a $\langle v_1 - v_2, x\rangle = 0$. En particulier,

$$\langle v_1 - v_2, v_1 - v_2\rangle = 0.$$

Comme $\langle\ ,\ \rangle$ est un produit scalaire, cela implique que $v_1 - v_2 = 0$, soit encore $v_1 = v_2$. Cela achève la démonstration. □

5.13. Corollaire. *Soit E un espace euclidien orienté de dimension $n \geqslant 1$. Pour tous $x_1, \ldots, x_{n-1} \in E$, il existe un unique vecteur $x_1 \wedge \cdots \wedge x_{n-1} \in E$ tel que*

$$[x_1, \ldots, x_{n-1}, x] = \langle x_1 \wedge \cdots \wedge x_{n-1}, x\rangle, \qquad \textit{pour tout } x \in E.$$

Si $\mathbf{e} = (e_1, \ldots, e_n)$ est une base orthonormée directe de E, et si M est la matrice dont les colonnes sont $[x_1]_{\mathbf{e}}, \ldots, [x_{n-1}]_{\mathbf{e}}$, alors on a

$$x_1 \wedge \cdots \wedge x_{n-1} = \sum_{i=1}^{n} (-1)^{n+i} \det(M_i) e_i,$$

où $M_i \in \mathrm{M}_{n-1}(\mathbb{R})$ est la matrice extraite de M obtenue en supprimant la ligne i.

Démonstration. Soient $x_1, \ldots, x_{n-1} \in E$. Puisque le produit mixte est en particulier n-linéaire, l'application

$$\begin{aligned} \varphi : E &\longrightarrow K \\ x &\longmapsto [x_1, \ldots, x_{n-1}, x] \end{aligned}$$

est linéaire.

De plus, si $\mathbf{e} = (e_1, \ldots, e_n)$ est une base orthonormée directe de E, alors pour tout $i \in [\![1, n]\!]$, on a

$$\varphi(e_i) = [x_1, \ldots, x_{n-1}, e_i] = \det_{\mathbf{e}}(x_1, \ldots, x_{n-1}, e_i).$$

En utilisant la définition du déterminant et en développant par rapport à la dernière colonne, on en déduit que $\varphi(e_i) = (-1)^{n+i} \det(M_i)$. Il suffit alors d'appliquer le lemme précédent pour conclure. □

5.14. Définition. Soit E un espace euclidien orienté de dimension $n \geqslant 1$, et soient $x_1, \ldots, x_{n-1} \in E$. Le vecteur $x_1 \wedge \cdots \wedge x_{n-1} \in E$ s'appelle le *produit vectoriel* de $x_1, \ldots, x_{n-1}$.

5.15. Exemple. Soit $E = \mathbb{R}^3$, que l'on munit de son produit scalaire usuel, et que l'on oriente grâce à sa base canonique. Si $x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$ et $y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$, on a

$$x \wedge y = \begin{pmatrix} x_2y_3 - x_3y_2 \\ x_3y_1 - x_1y_3 \\ x_1y_2 - x_2y_1 \end{pmatrix}.$$

On retrouve alors le produit vectoriel usuel.

5.16. Proposition. *Soit $(E, \langle\ ,\ \rangle)$ un espace euclidien orienté de dimension finie $n \geqslant 1$. L'application*

$$\begin{aligned} E \times \cdots \times E &\longrightarrow \mathbb{R} \\ (x_1, \ldots, x_{n-1}) &\longmapsto x_1 \wedge \cdots \wedge x_{n-1} \end{aligned}$$

est $(n-1)$-linéaire alternée.

De plus, pour tous $x_1, \ldots, x_{n-1} \in E$, on a les propriétés suivantes :

(1) *le vecteur $x_1 \wedge \cdots \wedge x_{n-1}$ est orthogonal à $x_1, \ldots, x_{n-1}$;*

(2) *on a $x_1 \wedge \cdots \wedge x_{n-1} \neq 0$ si, et seulement si, la famille $(x_1, \ldots, x_{n-1})$ est libre.*

Dans ce cas, la famille

$$(x_1, \ldots, x_{n-1}, x_1 \wedge \cdots \wedge x_{n-1})$$

est une base directe de E ;

(3) *si $(x_1, \ldots, x_{n-1})$ est une famille orthonormée, la famille*

$$(x_1, \ldots, x_{n-1}, x_1 \wedge \cdots \wedge x_{n-1})$$

est une base orthonormée directe de E.

Démonstration. Le fait que l'application

$$\begin{aligned} E \times \cdots \times E &\longrightarrow \mathbb{R} \\ (x_1, \ldots, x_{n-1}) &\longmapsto x_1 \wedge \cdots \wedge x_{n-1} \end{aligned}$$

soit $(n-1)$-linéaire alternée découle par exemple de la formule explicite donnée dans le corollaire 5.13 et des propriétés du déterminant.

Démontrons (1). Par définition du produit vectoriel, pour tout $i \in [\![1, n]\!]$, on a

$$\langle x_1 \wedge \cdots \wedge x_{n-1}, x_i \rangle = [x_1, \ldots, x_{n-1}, x_i] = 0,$$

car le produit mixte est une application alternée, d'où le résultat.

Soient $\mathbf{e}$ est une base orthonormée directe de E, soient $x_1, \ldots, x_{n-1} \in E$, et soit M la matrice dont les colonnes sont $[x_1]_{\mathbf{e}}, \ldots, [x_{n-1}]_{\mathbf{e}}$.

Toujours d'après la formule explicite mentionnée plus haut, $x_1 \wedge \cdots \wedge x_{n-1}$ est non nul si, et seulement si, la matrice M définie possède une sous-matrice carrée de taille $n-1$ de déterminant non nul. Cela revient à dire que M est de rang $n-1$, ou encore que la famille $(x_1, \ldots, x_{n-1})$ est libre.

Dans ce cas, on a alors

$$[x_1, \ldots, x_{n-1}, x_1 \wedge \cdots \wedge x_{n-1}] = \langle x_1 \wedge \cdots \wedge x_{n-1}, x_1 \wedge \cdots \wedge x_{n-1} \rangle > 0.$$

Par le lemme 5.10 (2), on obtient que la famille

$$(x_1, \ldots, x_{n-1}, x_1 \wedge \cdots \wedge x_{n-1})$$

est une base directe, d'où (2).

Démontrons (3). Pour cela, on complète $(x_1, \ldots, x_{n-1})$ en une base orthonormée $\mathbf{x} = (x_1, \ldots, x_n)$. Quitte à remplacer x_n par $-x_n$, on peut supposer que $\mathbf{x}$ est une base orthonormée directe. Par le point (1), $x_1 \wedge \cdots \wedge x_{n-1}$ est dans $\mathrm{Vect}_{\mathbb{R}}(x_1, \ldots, x_{n-1})^{\perp} = \mathbb{R}x_n$. Il existe donc $\lambda \in \mathbb{R}$ tel que

$$x_1 \wedge \cdots \wedge x_{n-1} = \lambda x_n.$$

Par ailleurs, puisque $\mathbf{x}$ est une base orthonormée directe, on a

$$[x_1, \ldots, x_n] = \det_{\mathbf{x}}(\mathbf{x}) = 1,$$

par définition du produit mixte. D'après la définition du produit vectoriel, on a alors

$$\begin{aligned} \langle x_1 \wedge \cdots \wedge x_{n-1}, x_1 \wedge \cdots \wedge x_{n-1} \rangle &= [x_1, \ldots, x_{n-1}, x_1 \wedge \cdots \wedge x_{n-1}] \\ &= [x_1, \ldots, x_{n-1}, \lambda x_n] \\ &= \lambda [x_1, \ldots, x_{n-1}, x_n] \\ &= \lambda. \end{aligned}$$

Mais l'on a aussi

$$\langle x_1 \wedge \cdots \wedge x_{n-1}, x_1 \wedge \cdots \wedge x_{n-1} \rangle = \lambda^2 \langle x_n, x_n \rangle = \lambda^2,$$

d'où $\lambda^2 = \lambda$. Comme $(x_1, \ldots, x_{n-1})$ est libre, $x_1 \wedge \cdots \wedge x_{n-1} \neq 0$ par (2), et par suite $\lambda \neq 0$. Par conséquent, $\lambda = 1$. On a ainsi $x_n = x_1 \wedge \cdots \wedge x_{n-1}$, et donc $(x_1, \ldots, x_{n-1}, x_1 \wedge \cdots \wedge x_{n-1}) = \mathbf{x}$ est une base orthonormée directe. Cela achève la démonstration. □

5.17. Corollaire. *Soit E un espace euclidien orienté de dimension $n \geqslant 1$, et soit H un hyperplan de E, de base $(x_1, \ldots, x_{n-1})$. Alors :*

(1) *la droite $D = H^{\perp}$ est engendrée par $x_1 \wedge \cdots \wedge x_{n-1}$;*

(2) *on a $H = \{x \in E \mid \langle x_1 \wedge \cdots \wedge x_{n-1}, x \rangle = 0\}$;*

(3) *pour tout $x \in E$, la projection orthogonale $p_H(x)$ de x sur H est donnée par*

$$p_H(x) = x - \frac{\langle x_1 \wedge \cdots \wedge x_{n-1}, x \rangle}{\langle x_1 \wedge \cdots \wedge x_{n-1}, x_1 \wedge \cdots \wedge x_{n-1} \rangle} x_1 \wedge \cdots \wedge x_{n-1} ;$$

(4) *pour tout $x \in E$, la distance $d(x,H)$ de x à H est donnée par*

$$d(x,H) = \frac{|\langle x_1 \wedge \cdots \wedge x_{n-1}, x\rangle|}{\sqrt{\langle x_1 \wedge \cdots \wedge x_{n-1}, x_1 \wedge \cdots \wedge x_{n-1}\rangle}}.$$

Démonstration. D'après la proposition 5.16 (1), la droite D contient le vecteur $x_1 \wedge \cdots \wedge x_{n-1}$. Comme $(x_1, \ldots, x_{n-1})$ est une famille libre, ce vecteur est non nul par le point (2) de cette même proposition, et engendre donc la droite D. Cela démontre (1). Puisque $H = (H^\perp)^\perp = D^\perp$, l'alinéa (2) en découle immédiatement.

Si $x \in E$, par définition de $p_H(x)$, on a $x - p_H(x) \in D$, et donc

$$x = p_H(x) + \lambda x_1 \wedge \cdots \wedge x_{n-1},$$

avec $\lambda \in \mathbb{R}$. Il reste donc à calculer λ. Comme $p_H(x) \in H$, (2) nous donne alors

$$\langle x_1 \wedge \cdots \wedge x_{n-1}, x\rangle = \lambda \langle x_1 \wedge \cdots \wedge x_{n-1}, x_1 \wedge \cdots \wedge x_{n-1}\rangle,$$

d'où (3). Enfin, d'après les propriétés bien connues de la projection orthogonale, on a

$$d(x,H) = \sqrt{\langle x - p_H(x), x - p_H(x)\rangle}.$$

Le point (4) découle alors immédiatement du point (3). □

6. Exercices

E1. Dans les cas suivants, déterminer l'image du vecteur $v \in E$ par la symétrie $s \in \mathscr{L}(E)$ par rapport à F parallèlement à G.

a. $E = \mathbb{R}^2$, F est l'axe Ox, G est l'axe Oy, et $v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}$.

b. $E = \mathbb{R}^3$, F est le plan Oxy, G est l'axe Oz, et $v = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$.

c. $E = \mathbb{R}^3$, F est l'axe Oz, G est le plan Oxy, et $v = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$.

d. $E = \mathbb{R}^3$, F est le plan Oxy, $G = \mathbb{R}\begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}$, et $v = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}$.

E2. Soit E un espace euclidien, et soient $x_1, \ldots, x_p \in E$. On pose

$$G(x_1, \ldots, x_p) = (\langle x_i, x_j\rangle)_{i,j} \in \mathrm{M}_p(\mathbb{R}).$$

Soit $F = \mathrm{Vect}_{\mathbb{R}}(x_1, \ldots, x_p)$.

a. Soit $\mathbf{e} = (e_1, \ldots, e_r)$ une base orthonormée de F, et soit $M \in \mathrm{M}_{r\times p}(\mathbb{R})$ la matrice dont les vecteurs colonnes sont $[x_1]_{\mathbf{e}}, \ldots, [x_p]_{\mathbf{e}}$.

Montrer que $G(x_1, \ldots, x_p) = M^t M$.

b. En déduire que M et $G(x_1, \ldots, x_p)$ ont même noyau, puis même rang.

Indication. Si $GX = 0$, remarquer que $X^t GX = 0$ et utiliser la question précédente.

c. Montrer que $(x_1, \ldots, x_p)$ est liée si, et seulement si, on a

$$\det\big(G(x_1, \ldots, x_p)\big) = 0.$$

Montrer également que si $(x_1, \ldots, x_p)$ est libre, alors

$$\det\big(G(x_1, \ldots, x_p)\big) > 0.$$

d. On suppose que $(x_1, \ldots, x_p)$ est libre. Soit $x \in E$.

(1) Utiliser la question précédente pour montrer que

$$\det\big(G(x, x_1, \ldots, x_p)\big) = \det\big(G(x - p_F(x), x_1, \ldots, x_p)\big).$$

(2) Développer $\det\big(G(x - p_F(x), x_1, \ldots, x_p)\big)$ par rapport à la première colonne, et en déduire la formule

$$d(x, F) = \sqrt{\frac{\det\big(G(x, x_1, \ldots, x_p)\big)}{\det\big(G(x_1, \ldots, x_p)\big)}}.$$

E3. Soit E un ensemble, et soit X une partie de E. On dit qu'une application $u : E \longrightarrow E$ laisse X globalement invariant lorsque $f(X) = X$.

Dans les cas suivants, déterminer l'ensemble des endomorphismes de $\mathbb{R}^2$ qui laissent X globalement invariant. On caractérisera par exemple ses éléments par leurs matrices représentatives. Dire lesquelles sont des isométries, et préciser leurs éléments caractéristiques.

a. X est le cercle-unité.

b. $X = \mathbb{Z}^2$.

c. X est formé des deux points $(1, 0)$ et $(-1, 0)$.

d. X est l'ensemble des sommets du triangle équilatéral

$$A = (1, 0), \quad B = \Big(\cos\big(\tfrac{2\pi}{3}\big), \sin\big(\tfrac{2\pi}{3}\big)\Big), \quad C = \Big(\cos\big(\tfrac{4\pi}{3}\big), \sin\big(\tfrac{4\pi}{3}\big)\Big).$$

e. X est l'ensemble des sommets du carré de côté 2, centré en O.

f. X est l'ensemble des sommets du rectangle

$$A = (2,1), \quad B = (-2,1), \quad C = (-2,-1), \quad D = (2,-1).$$

Indication. On pourra d'abord remarquer (et justifier !) que u est nécessairement bijective dans chaque cas.

E4. On considère dans $\mathbb{R}^3$ l'ensemble des points (x_1, x_2, x_3), $x_i = \pm 1$ formant les sommets d'un cube C. Combien y a-t-il d'isométries vectorielles de $\mathbb{R}^3$ laissant C globalement invariant ?

E5. Soit E un espace euclidien, et soit $u \in \mathscr{L}(E)$. Montrer que les propositions suivantes sont équivalentes :

(1) u est une symétrie orthogonale ;

(2) il existe une base orthonormée $\mathbf{e}$ telle que $\mathrm{Mat}(u\,;\mathbf{e})$ est symétrique ;

(3) pour toute base orthonormée $\mathbf{e}$, $\mathrm{Mat}(u\,;\mathbf{e})$ est symétrique.

E6. Déterminer toutes les matrices triangulaires supérieures de $\mathrm{M}_n(\mathbb{R})$ qui commutent avec leur transposée.

E7. Soit E un espace euclidien orienté de dimension 3. Soit $u \in \mathrm{O}(E)$. On pose $E_1 = \mathrm{Ker}(u - \mathrm{Id}_E)$ et $E_{-1} = \mathrm{Ker}(u + \mathrm{Id}_E)$. Soit $\mathbf{e}$ une base orthonormée directe de E fixée, et soit $A = \mathrm{Mat}(u\,;\mathbf{e})$.

a. On suppose que A est symétrique. Justifier que u est la symétrie orthogonale par rapport à E_1 parallèlement à E_{-1}.

b. On suppose que A n'est pas symétrique. Soit $\varepsilon \in \{1, -1\}$.

(i) Si $\det(u) = \varepsilon$, montrer que $\dim(E_\varepsilon) = 1$, et que la restriction de l'endomorphisme u à $F = E_\varepsilon^\perp$ est une rotation plane.

(ii) Soit v un vecteur unitaire dirigeant E_ε. Montrer que l'angle θ de la rotation u_F est entièrement déterminé (en mesure, à un multiple de 2π près) par les deux conditions suivantes :

▷ $\mathrm{Tr}(u) = \varepsilon + 2\cos(\theta)$;

▷ $\sin(\theta)$ est du signe de $[v, w, u(w)]$, où w est n'importe quel vecteur non colinéaire à v.

c. Soit $E = \mathbb{R}^3$, muni de son produit scalaire usuel et orienté par sa base canonique. Soit $A = \dfrac{1}{15}\begin{pmatrix} -11 & -10 & 2 \\ 2 & -5 & -14 \\ -10 & 10 & -5 \end{pmatrix}$.

Montrer que $A \in \mathrm{O}_3(\mathbb{R})$. Déterminer si c'est une symétrie orthogonale, une rotation, ou la composée d'une rotation et d'une réflexion orthogonale. Calculer les éléments caractéristiques (selon les cas : axe, angle et plan de la rotation, plan de la réflexion, sous-espace par rapport auquel on fait la symétrie, etc.).

E8. Dans cet exercice, on suppose le lecteur familier avec la notion de connexité et de connexité par arcs.

a. Soit E un espace hermitien. Montrer que $\mathrm{U}(E)$ est connexe par arcs.

b. Soit E un espace euclidien. Montrer que $\mathrm{O}^+(E)$ est connexe par arcs. Combien de composantes connexes le groupe $\mathrm{O}(E)$ possède-t-il ?

Indication. Se ramener aux cas des groupes $\mathrm{U}_n(\mathbb{C})$ et $\mathrm{O}_n^+(\mathbb{R})$. Montrer qu'il suffit alors de trouver un chemin continu qui relie n'importe quel élément de ces groupes à la matrice identité.

E9. Soit E un espace euclidien orienté de dimension $n \geqslant 2$. Soient $a, b \in E$ avec a non nul. Résoudre l'équation

$$a \wedge x = b, \quad x \in E.$$

E10. Soit E un espace euclidien orienté de dimension $n \geqslant 2$.

a. Soient $a_1, \ldots, a_{n-2} \in E$. Montrer que l'endomorphisme

$$\begin{aligned} E &\longrightarrow E \\ x &\longmapsto a_1 \wedge \cdots \wedge a_{n-2} \wedge x \end{aligned}$$

est antisymétrique.

b. On suppose que $n = 3$. Montrer qu'un endomorphisme $u \in \mathscr{L}(E)$ est antisymétrique si, et seulement si, il existe $a \in E$ tel que

$$u(x) = a \wedge x, \quad \text{pour tout } x \in E.$$

Montrer que dans ce cas, a est unique.

Deuxième partie

Groupes : il faut agir !!

Chapitre VI

Propriétés élémentaires des groupes

Dans ce chapitre, nous introduisons les définitions et résultats élémentaires de la théorie des groupes.

1. Généralités

Bien que nous ayons déjà introduit cette notion dans le chapitre I (cf. définition I-3.9), rappelons la définition d'un groupe.

1.1. Définition. Un *groupe* est un couple $(G, *)$, où G est un ensemble non vide et $* : G \times G \longrightarrow G$ est une loi interne vérifiant les propriétés suivantes :

(1) la loi $*$ est associative : pour tous x, y, $z \in G$, on a $(x*y)*z = x*(y*z)$;

(2) la loi $*$ possède un élément neutre : il existe un élément $e \in G$ tel que l'on ait $x * e = e * x = x$, pour tout $x \in G$;

(3) tout élément possède un symétrique : pour tout $x \in G$, il existe $x' \in G$ tel que $x * x' = x' * x = e$.

D'après les lemmes I-3.5 et I-3.8, l'élément neutre est unique, et un symétrique de $x \in G$ est unique. On le note x^{-1}, et on l'appelle *l'inverse* de x.

Un groupe $(G, *)$ est dit *abélien* (ou commutatif) si l'on a de plus

$$x * y = y * x, \qquad \text{pour tous } x, y \in G.$$

Passons aux exemples de groupes.

1.2. Exemples

(1) Un singleton $\{e\}$ peut être muni trivialement d'une structure de groupe en posant $e * e = e$.

(2) Les couples $(\mathbb{Z}, +)$, $(\mathbb{R}, +)$, $(\mathbb{R}^\times, \cdot)$, $(\mathbb{C}, +)$, $(\mathbb{C}^\times, \cdot)$ sont des groupes abéliens.

(3) Le couple $\big(\mathrm{GL}_n(\mathbb{C}), \cdot\big)$ est un groupe, non abélien si $n \geqslant 2$.

(4) Soit E un K-espace vectoriel, où K est un corps. Alors, l'ensemble des applications linéaires bijectives $f : E \longrightarrow E$ est un groupe, pour la composition des applications, noté $\mathrm{GL}(E)$.

(5) Si E est un K-espace vectoriel, $(E, +)$ est un groupe abélien, comme par exemple $\big(\mathrm{M}_n(\mathbb{C}), +\big)$ et $(\mathbb{C}^n, +)$.

(6) Le couple $(\mathbb{Z}, \cdot)$ n'est pas un groupe : par exemple, 2 n'a pas d'inverse.

(7) Soit E un ensemble. L'ensemble $\mathfrak{S}(E)$ des bijections de E dans lui-même, muni de la composition, est un groupe, appelé *groupe des permutations* de E ou *groupe symétrique* sur E. Il est non abélien dès que $|E| \geqslant 3$. Si $E = [\![1, n]\!]$, on obtient le groupe $\mathfrak{S}_n$ défini dans le paragraphe IV-1.

(8) Si $(G, *)$ est un groupe et X est un ensemble, on note G^X l'ensemble des applications $f : X \longrightarrow G$. Pour toutes applications $f, g \in G^X$, on définit l'application $f * g : X \longrightarrow G$ par

$$(f * g)(x) = f(x) * g(x), \qquad \text{pour tout } x \in X.$$

Alors, $(G^X, *)$ est un groupe. (Quels sont le neutre et l'inverse d'un élément ?)

(9) Si $(G, *)$ est un groupe, l'ensemble des suites $(g_n)_{n \geqslant 0}$ d'éléments de G est un groupe. C'est un cas particulier du cas précédent, puisqu'une suite n'est rien d'autre qu'une application $f : \mathbb{N} \longrightarrow G$.

(10) Si $(G, *)$ et $(G', \bullet)$ sont deux groupes, l'ensemble $G \times G'$, muni de la loi

$$\begin{aligned} (G \times G') \times (G \times G') &\longrightarrow G \times G' \\ \big((g_1, g_1'), (g_2, g_2')\big) &\longmapsto (g_1 * g_2, g_1' \bullet g_2') \end{aligned}$$

est un groupe. Le groupe obtenu s'appelle le *produit direct (externe)* de G et G' .

Par exemple, $(\mathbb{R}^2, +)$ est le produit direct $(\mathbb{R}, +) \times (\mathbb{R}, +)$.

Nous laissons le soin au lecteur de détailler les exemples.

Notations

(1) Dans les énoncés, on omettra de préciser la loi $*$, et l'on parlera du groupe G, plutôt que du groupe $(G, *)$.

(2) Le produit de deux éléments x et y de G est en général noté xy, et le neutre est noté 1_G ou 1 s'il n'y a pas d'ambiguïté.

Attention néanmoins, cette loi n'est pas forcément la multiplication au sens usuel ; c'est juste une notation commode lorsque l'on travaille avec une structure de groupe non spécifique. Lorsque G est abélien, on utilise aussi la notation additive « + ». Dans ce cas précis, le neutre est noté 0_G ou 0, et l'inverse (le symétrique) de x est noté $-x$.

(3) Si $x_1, \ldots, x_n \in G$, le produit $x_1 \cdots x_n$, noté aussi $\prod_{i=1}^{n} x_i$, est défini par récurrence de la manière suivante :

(1) si $n = 0$, $x_1 \cdots x_n = 1_G$;

(2) si $n \geqslant 1$, on pose $x_1 \cdots x_n = (x_1 \cdots x_{n-1})x_n$.

Remarquons que par associativité, pour tous k, $n \geqslant 0$, on a

$$(x_1 \cdots x_k)(x_{k+1} \cdots x_n) = x_1 \cdots x_n.$$

(4) Si $n \in \mathbb{Z}$, on pose $x^n = x \cdots x$ si $n \geqslant 0$ et $x^n = x^{-1} \cdots x^{-1}$ si $n < 0$. (En notation additive, on le note $n \cdot x$.)

On a alors

$$x^{n+m} = x^n x^m \quad \text{et} \quad (x^n)^m = x^{nm}, \qquad \text{pour tous } n, m \in \mathbb{Z}.$$

Ces deux égalités sont faciles, mais fastidieuses à démontrer. Leur vérification est lâchement laissée au lecteur.

Attention ! En général, $(xy)^n \neq x^n y^n$ si G n'est pas abélien.

1.3. Lemme *Soit G un groupe. Alors :*

(1) *la loi de groupe est simplifiable : pour tous x, y, $z \in G$, on a*

$$xz = yz \implies x = y \quad \text{et} \quad zx = zy \implies x = y$$

(2) *pour tous $x, x_1, \ldots, x_n \in G$, on a*

$$(x^{-1})^{-1} = x \quad \text{et} \quad (x_1 \cdots x_n)^{-1} = x_n^{-1} \cdots x_1^{-1}$$

(3) *pour tout $a \in G$, les translations à gauche et à droite*

$$\ell_a : \begin{array}{ccc} G & \longrightarrow & G \\ x & \longmapsto & ax \end{array} \quad \text{et} \quad r_a : \begin{array}{ccc} G & \longrightarrow & G \\ x & \longmapsto & xa \end{array}$$

sont bijectives.

Démonstration

(1) Il suffit de multiplier par l'inverse et d'utiliser l'associativité.

(2) Puisque $x^{-1}x = xx^{-1} = 1_G$, x est l'unique inverse de x^{-1}, d'où le premier point. Pour le second, il suffit de traiter le cas $n = 2$ et de conclure par récurrence. On a

$$x_1x_2(x_2^{-1}x_1^{-1}) = x_1(x_2x_2^{-1})x_1^{-1} = x_1 1_G x_1^{-1} = x_1x_1^{-1} = 1_G,$$

et de même $(x_2^{-1}x_1^{-1})x_1x_2 = 1_G$.

(3) On vérifie que les applications inverses sont $\ell_{a^{-1}}$ et $r_{a^{-1}}$ respectivement. □

Notation. Si E est un ensemble fini, on notera $|E|$ son nombre d'éléments.

1.4. Définition. Si G est un groupe fini, son nombre d'éléments $|G|$ est appelé *l'ordre* de G.

1.5. Exemples

(1) Si E est un ensemble à n éléments, $\mathfrak{S}(E)$ est un groupe fini d'ordre $n!$.

(2) Si $n \geqslant 1$, le groupe additif $\mathbb{Z}/n\mathbb{Z}$ (introduit dans le paragraphe II-6) est un groupe d'ordre n, d'après le lemme II-6.5.

Maintenant que les objets à étudier ont été définis, il nous faut définir les « flèches » entre les objets, appelées morphismes. Évidemment, de tels morphismes n'ont d'intérêt que s'ils « respectent » la structure de groupe.

1.6. Définition. Soient G, G' deux groupes.

Une application $f : G \longrightarrow G'$ est un *morphisme de groupes* si

$$f(xy) = f(x)f(y), \quad \text{pour tous } x, y \in G.$$

Attention ! La loi de groupe à droite est la loi de groupe dans G'. Il se peut très bien que G soit noté multiplicativement, et G' additivement, ou l'inverse.

Si $f : G_1 \longrightarrow G_2$ et $g : G_2 \longrightarrow G_3$ sont des morphismes de groupes, il en est de même de $g \circ f : G_1 \longrightarrow G_3$.

1.7. Exemples

L'application ci-après est un morphisme de groupes :

(1) $h_x : \mathbb{Z} \longrightarrow G$, $m \longmapsto x^m$, où G est un groupe quelconque et $x \in G$;

Il en est de même des applications qui suivent.

(2) $\begin{array}{rcl} f : \mathbb{Z} & \longrightarrow & \mathbb{Z} \\ n & \longmapsto & 2n \end{array}$;

(3) $\begin{array}{rcl} f : \mathbb{R}^\times & \longrightarrow & \mathbb{R}^\times \\ x & \longmapsto & x^3 \end{array}$;

(4) $\exp : \mathbb{R} \longrightarrow \mathbb{R}^\times$;

(5) $\ln : \mathbb{R}_+^\times \longrightarrow \mathbb{R}$;

(6) $\det : \mathrm{GL}_2(\mathbb{C}) \longrightarrow \mathbb{C}^\times$.
En revanche, $\det : \mathrm{M}_2(\mathbb{R}) \longrightarrow \mathbb{R}$ n'est **pas** un morphisme de groupes ;

(7) plus généralement, si K est un corps et E est un K-espace vectoriel de dimension finie, l'application $\det : \mathrm{GL}(E) \longrightarrow K^\times$ est un morphisme de groupes multiplicatifs ;

(8) pour tout $n \geqslant 2$, la signature $\varepsilon : \mathfrak{S}_n \longrightarrow \{\pm 1\}$ est un morphisme de groupes.

Encore une fois, on laisse au lecteur le soin de détailler les exemples.

La remarque suivante est très pratique et sera utilisée sans référence ultérieure.

1.8. Remarque. Soit $f : G \longrightarrow G'$ un morphisme de groupes. Alors :

(1) on a $f(1_G) = 1_{G'}$;

(2) pour tout $x \in G$, on a $f(x^{-1}) = f(x)^{-1}$.

En effet, on a

$$f(1_G) = f(1_G \cdot 1_G) = f(1_G) f(1_G),$$

et l'on conclut par simplification dans G'. On a aussi

$$f(x) f(x^{-1}) = f(x x^{-1}) = f(1_G) = 1_{G'},$$

et de même $f(x^{-1}) f(x) = 1_{G'}$, donc $f(x^{-1})$ est l'inverse de $f(x)$.

1.9. Définition. Un *isomorphisme* de groupes est un morphisme de groupes bijectif.

Deux groupes G et G' sont *isomorphes* s'il existe au moins un isomorphisme de groupes $f : G \longrightarrow G'$. On note cela $G \simeq G'$.

Un *automorphisme* d'un groupe G est un isomorphisme de groupes de G sur lui-même. L'ensemble des automorphismes de G est noté $\mathrm{Aut}(G)$.

1.10. Remarques

(1) Pour tout groupe G, l'identité $\mathrm{Id}_G : G \longrightarrow G$ est un isomorphisme. On a donc $G \simeq G$, pour tout groupe G.

(2) Si $f : G \longrightarrow G'$ est un isomorphisme, alors $f^{-1} : G' \longrightarrow G$ est aussi un morphisme de groupes, et est donc un isomorphisme de groupes.
En effet, soient x', $y' \in G'$. On a alors

$$f\big(f^{-1}(x')f^{-1}(y')\big) = f\big(f^{-1}(x')\big)f\big(f^{-1}(y')\big) = x'y',$$

et donc $f^{-1}(x'y') = f^{-1}(x')f^{-1}(y')$ par définition de f^{-1}.
En particulier, si $G \simeq G'$, alors $G' \simeq G$.

(3) Si $f_1 : G_1 \longrightarrow G_2$ et $f_2 : G_2 \longrightarrow G_3$ sont des isomorphismes de groupes, il en est de même de $f_2 \circ f_1 : G_1 \longrightarrow G_3$.
Par conséquent, si $G_1 \simeq G_2$ et $G_2 \simeq G_3$, alors $G_1 \simeq G_3$.
La vérification est laissée au lecteur.

(4) Les trois points précédents entraînent aisément que l'ensemble $\mathrm{Aut}(G)$ est un groupe pour la composition des applications.

1.11. Exemples. Les applications suivantes sont des isomorphismes de groupes :

(1) $\begin{array}{rcl} f : \mathbb{R}^\times & \longrightarrow & \mathbb{R}^\times \\ x & \longmapsto & x^3 \end{array}$;

(2) $\exp : \mathbb{R} \longrightarrow \mathbb{R}_+^\times$;

(3) $\ln : \mathbb{R}_+^\times \longrightarrow \mathbb{R}$;

(4) si E est un K-espace vectoriel de dimension n, le choix d'une base $\mathbf{e}$ induit un isomorphisme de groupes

$$\begin{array}{rcl} \mathrm{GL}(E) & \xrightarrow{\sim} & \mathrm{GL}_n(K) \\ u & \longmapsto & \mathrm{Mat}(u\,;\mathbf{e}). \end{array}$$

Le lemme suivant donne toute une famille d'exemples d'automorphismes de groupes.

1.12. Lemme. *Soit G un groupe, et soit $g \in G$. L'application*

$$\begin{array}{rcl} \mathrm{Int}_g : G & \longrightarrow & G \\ x & \longmapsto & gxg^{-1} \end{array}$$

est un automorphisme de G, appelé la conjugaison par g.

De plus, l'application

$$\begin{array}{rcl} \mathrm{Int} : G & \longrightarrow & \mathrm{Aut}(G) \\ g & \longmapsto & \mathrm{Int}_g \end{array}$$

est un morphisme de groupes.

Démonstration. Pour tous x, $y \in G$, on a

$$\mathrm{Int}_g(xy) = gxyg^{-1} = gxg^{-1}gyg^{-1} = \mathrm{Int}_g(x)\mathrm{Int}_g(y).$$

Ainsi, Int_g est un morphisme de groupes. De plus, pour tout $x \in G$, et pour tous $g, g' \in G$, on a

$$\mathrm{Int}_{gg'}(x) = gg'x(gg')^{-1} = gg'xg'^{-1}g^{-1} = g\mathrm{Int}_{g'}(x)g^{-1} = \mathrm{Int}_g\big(\mathrm{Int}_{g'}(x)\big).$$

Par conséquent, on a

$$\mathrm{Int}_g \circ \mathrm{Int}_{g'} = \mathrm{Int}_{gg'}.$$

On a en particulier

$$\mathrm{Int}_g \circ \mathrm{Int}_{g^{-1}} = \mathrm{Int}_{g^{-1}} \circ \mathrm{Int}_g = \mathrm{Int}_{1_G} = \mathrm{Id}_G,$$

et donc Int_g est bijectif, d'inverse $\mathrm{Int}_{g^{-1}}$. Le dernier point est clair, au vu des calculs précédents. □

1.13. Définition. Un *automorphisme intérieur* d'un groupe G est un automorphisme de la forme $\mathrm{Int}(g)$, $g \in G$. L'ensemble $\mathrm{Int}(G)$ des automorphismes intérieurs est un groupe pour la composition des applications.

Les automorphismes intérieurs jouent un rôle important en théorie des groupes (cf. la notion de sous-groupe distingué dans le paragraphe suivant). Nous finissons ce paragraphe en introduisant quelques notations et définitions.

Notation. Si P est une partie de G, pour tout $g \in G$, on note gPg^{-1} l'image de P par Int_g. Autrement dit, on a

$$gPg^{-1} = \{gxg^{-1} \mid x \in P\}.$$

Puisque Int_g est un automorphisme, gPg^{-1} est fini si, et seulement si, P est fini, et dans ce cas, on a

$$|gPg^{-1}| = |P|,$$

et cela pour tout $g \in G$.

1.14. Définition. Soit G un groupe, et soient deux parties P, P' de G. On dit que P et P' sont *conjuguées* dans G s'il existe $g \in G$ tel que $P' = gPg^{-1}$.

Deux éléments x et x' de G sont dits *conjugués* dans G s'il existe $g \in G$ (non nécessairement unique) tel que $x' = gxg^{-1}$. La *classe de conjugaison* de x dans G est l'ensemble $\mathrm{Conj}_G(x)$ des éléments de G conjugués à x, c'est-à-dire l'ensemble

$$\mathrm{Conj}_G(x) = \{gxg^{-1} \mid g \in G\}.$$

Le lecteur vérifiera sans peine que la relation « être conjugués dans G » est une relation d'équivalence sur G, dont les classes d'équivalence sont les classes de conjugaison d'éléments de G.

1.15. Remarque. Déterminer toutes les classes de conjugaison d'un groupe est une question importante en théorie des groupes, mais difficile à résoudre. Par exemple, si $G = \mathrm{GL}_n(K)$, cela revient à savoir décider si deux matrices données M et M' sont semblables ou non. Ce problème occupera une bonne partie des chapitres de ce livre consacrés à l'algèbre linéaire.

Il y a quand même quelques groupes où ce problème peut être résolu facilement. C'est le cas du groupe symétrique, comme nous le verrons dans le chapitre VIII (cf. corollaire VIII-2.25), ou dans le cas du groupe alterné avec un peu plus d'efforts (cf. exercice VIII-33).

Il y a aussi des cas où il est possible de déterminer facilement les classes de conjugaison de certains éléments, ce qui peut déjà grandement aider dans l'étude d'un groupe donné, comme nous le verrons au fur et à mesure.

1.16. Exemples

(1) Si G est un groupe, et si $x \in G$, on a $\mathrm{Conj}_G(x) = \{x\}$ si, et seulement si, x commute avec tout élément de G.

En particulier, si G est abélien, on a $\mathrm{Conj}_G(x) = \{x\}$, pour tout $x \in G$.

(2) Si E est un K-espace vectoriel de dimension $n \geqslant 2$, toutes les transvections sont conjuguées dans $\mathrm{GL}(E)$. Si de plus $n \geqslant 3$, elles sont aussi conjuguées dans $\mathrm{SL}(E)$ (cf. proposition IV-5.11).

(3) Soit E un espace euclidien de dimension $n \geqslant 1$. Alors, deux réflexions orthogonales sont conjuguées par un élément de $\mathrm{O}^+(E)$ d'après le théorème V-3.11. En particulier, les réflexions orthogonales sont conjuguées dans $\mathrm{O}(E)$.

De même, deux renversements sont conjugués dans $\mathrm{O}^+(E)$, d'après le théorème V-3.11.

2. Sous-groupes

Rappelons maintenant la notion de sous-groupe, déjà introduite dans le chapitre I (cf. définition I-3.12). C'est l'occasion pour nous d'en fournir quelques exemples et de mettre en évidence deux sous-groupes importants, qui viennent corrélés avec tout morphisme de groupes.

2.1. Définition. Soit G un groupe. Un *sous-groupe* de G est une partie H de G vérifiant les propriétés suivantes :

(1) on a $1_G \in H$;

(2) pour tous $x, y \in H$, $xy \in H$;

(3) pour tout $x \in H$, $x^{-1} \in H$.

Clairement, $\{1_G\}$ et G sont des sous-groupes de G. Un sous-groupe H sera dit *trivial* si $H = \{1_G\}$, *strict* si $H \neq G$, et *propre* s'il est distinct de $\{1_G\}$ et de G.

D'après le lemme I-3.14, un sous-groupe H de G est un groupe pour la loi induite sur H par celle de G. Pour montrer qu'un ensemble muni d'une loi est un groupe, il est ainsi plus rapide de montrer que c'est un sous-groupe d'un groupe connu, l'intérêt étant d'éviter de vérifier l'associativité.

De plus, le lemme suivant permet encore de raccourcir (parfois) les calculs.

2.2. Lemme. *Soit G un groupe, et soit H une partie de G. Alors, H est un sous-groupe de G si, et seulement si, H est non vide, et pour tous $x, y \in H$, on a $xy^{-1} \in H$.*

Démonstration. Si H vérifie les conditions du lemme, puisque H est non vide, il contient au moins un élément $x_0 \in H$, et contient donc $x_0 x_0^{-1} = 1_G$. Pour tout $x \in H$, on a alors $x^{-1} = 1_G x^{-1} \in H$. Enfin, si $x, y \in H$, alors d'après le point précédent, $y^{-1} \in H$, et donc $xy = x(y^{-1})^{-1} \in H$. Le sens direct étant clair, cela achève la démonstration. □

2.3. Exemples

(1) Il se peut fort bien qu'un groupe G n'ait aucun sous-groupe propre, comme $G = \{\pm 1\}$, par exemple.

(2) L'ensemble des fonctions continues de $\mathbb{R}$ dans $\mathbb{R}$ est un sous-groupe du groupe des fonctions de $\mathbb{R}$ dans $\mathbb{R}$ (pour l'addition des fonctions).

(3) L'ensemble des matrices triangulaires supérieures (resp. diagonales) est un sous-groupe de $\big(\mathrm{M}_n(\mathbb{C}), +\big)$.

(4) L'ensemble des matrices triangulaires supérieures inversibles est un sous-groupe de $\big(\mathrm{GL}_n(\mathbb{C}), \cdot\big)$.

(5) Si H est un sous-groupe de G et N est un sous-groupe de H, alors N est un sous-groupe de G.

(6) Si H_i est un sous-groupe de G_i, $i = 1, 2$, alors $H_1 \times H_2$ est un sous-groupe de $G_1 \times G_2$.

(7) La réunion de deux sous-groupes n'est pas un sous-groupe.
Par exemple, $H_1 = \{(0, y) \mid y \in \mathbb{R}\}$ et $H_2 = \{(x, 0) \mid x \in \mathbb{R}\}$ sont des sous-groupes de $\mathbb{R}^2$, mais pas $H_1 \cup H_2$.

Il est souvent ardu de décrire tous les sous-groupes d'un groupe G donné. Néanmoins, le cas de $\mathbb{Z}$ est plutôt aisé, et bien connu.

2.4. Proposition. *Les sous-groupes de $\mathbb{Z}$ sont tous de la forme*

$$n\mathbb{Z} = \{nm \mid m \in \mathbb{Z}\}, \quad n \in \mathbb{Z}.$$

Plus précisément, si H est un sous-groupe de $\mathbb{Z}$ non nul, alors on a $H = n\mathbb{Z}$, où n est le plus petit élément strictement positif de H.

Démonstration. Supposons tout d'abord $H = n\mathbb{Z}$. Alors, H est non vide, et puisque la différence de deux multiples de n est un multiple de n, on en déduit que H est un sous-groupe de $\mathbb{Z}$. Réciproquement, soit H un sous-groupe de $\mathbb{Z}$. Si $H = \{0\}$, le résultat est trivialement vrai. Supposons H non nul. Alors, l'ensemble $\{m \in H \mid m > 0\}$ est une partie non vide de $\mathbb{N}$, et admet donc un plus petit élément $n > 0$. Soit $m \in \mathbb{Z}$. Puisque $n \in H$ et que H est un sous-groupe, alors $nm = n + \cdots + n$ ou $(-n) + \cdots + (-n) \in H$, selon que m est positif ou négatif. On a donc l'inclusion $n\mathbb{Z} \subset H$. Réciproquement, soit $a \in H$. On peut écrire

$$a = qn + r, \quad q, r \in \mathbb{Z}, \quad \text{avec } r \in [\![0, n-1]\!].$$

Puisque a, $qn \in H$, on a $r = a - qn \in H$. Si l'on avait $r > 0$, cela contredirait le choix de n. Donc $r = 0$ et $a = qn \in n\mathbb{Z}$. Cela montre que $H \subset n\mathbb{Z}$, et donc $H = n\mathbb{Z}$. □

Le lemme suivant est utile pour construire de nouveaux sous-groupes.

2.5. Lemme. *Soit $f : G \longrightarrow G'$ un morphisme de groupes.*

(1) *Si H est un sous-groupe de G, alors $f(H)$ est un sous-groupe de G'.*

(2) *Si H' est un sous-groupe de G', alors $f^{-1}(H')$ est un sous-groupe de G.*

Démonstration. Soit H un sous-groupe de G, et soit

$$f(H) = \{f(h) \mid h \in H\}.$$

Puisque H est non vide, $f(H)$ aussi. Pour tout x, $y \in H$, on a

$$f(x)f(y)^{-1} = f(x)f(y^{-1}) = f(xy^{-1}).$$

Comme H est un sous-groupe, on a $xy^{-1} \in H$, et donc $f(x)f(y)^{-1} \in f(H)$. Ainsi, $f(H)$ est un sous-groupe de G'.

Si maintenant H' est un sous-groupe de G', soit

$$f^{-1}(H') = \{x \in G \mid f(x) \in H'\}.$$

Puisque $f(1_G) = 1_{G'} \in H'$, on a $1_G \in f^{-1}(H')$. En particulier, $f^{-1}(H')$ n'est pas vide. Soient maintenant $x, y \in f^{-1}(H')$. Alors,

$$f(xy^{-1}) = f(x)f(y)^{-1} \in H',$$

car $f(x), f(y) \in H'$ et H' est un sous-groupe. Ainsi, $xy^{-1} \in f^{-1}(H')$, d'où le résultat. □

2.6. Définition. Soit $f : G \longrightarrow G'$ un morphisme de groupes. Le *noyau* de f est le sous-groupe de G, noté $\mathrm{Ker}(f)$, défini par

$$\mathrm{Ker}(f) = f^{-1}(\{1_{G'}\}) = \{x \in G \mid f(x) = 1_{G'}\}.$$

L'image de f est le sous-groupe de G', noté $\mathrm{Im}(f)$, défini par

$$\mathrm{Im}(f) = f(G) = \{f(g) \mid g \in G\}.$$

Le lemme suivant donne une propriété très utile des morphismes de groupes.

2.7. Lemme. *Soit $f : G \longrightarrow G'$ un morphisme de groupes. Alors, f est injective si, et seulement si,* $\mathrm{Ker}(f) = \{1_G\}$.

Démonstration. Supposons f injective, et soit $x \in \mathrm{Ker}(f)$. On a alors

$$f(x) = 1_{G'} = f(1_G),$$

et par injectivité, on a $x = 1_G$. On a donc $\mathrm{Ker}(f) = \{1_G\}$. Inversement, supposons que $\mathrm{Ker}(f) = \{1_G\}$, et soient $x, y \in G$ tels que $f(x) = f(y)$. Alors, on a $f(x)f(y)^{-1} = f(xy^{-1}) = 1_{G'}$, et donc $xy^{-1} \in \mathrm{Ker}(f)$. Par hypothèse, $xy^{-1} = 1_G$, et donc $x = y$. Ainsi, f est injective. □

Nous allons maintenant définir une classe importante de sous-groupes.

2.8. Définition. Soit G un groupe, et soit H un sous-groupe. On dit que H est *distingué dans* G si l'on a

$$ghg^{-1} \in H, \quad \text{pour tout } h \in H \text{ et tout } g \in G.$$

Autrement dit, un sous-groupe H est distingué dans G s'il est stable par tout automorphisme intérieur de G. On dit aussi que H est un sous-groupe distingué de G. Nous utiliserons indifféremment les deux expressions dans cet ouvrage.

On note cela $H \triangleleft G$.

2.9. Exemples

(1) Si G est abélien, tout sous-groupe de G est distingué dans G.

(2) Les sous-groupes $\{1_G\}$ et G sont distingués dans G.

2.10. Remarque. Comme nous le verrons dans un chapitre ultérieur, les sous-groupes distingués interviendront de façon naturelle dans la construction des groupes quotients.

2.11. Définition. Soit G un groupe. Le *centre* d'un groupe est l'ensemble

$$Z(G) = \{z \in G \mid zg = gz, \quad \text{pour tout } g \in G\}.$$

C'est un sous-groupe abélien de G, distingué dans G.

Remarquons que $Z(G) = G$ si, et seulement si, G est abélien.

Donnons quelques exemples de calculs du centre d'un groupe.

2.12. Exemple. Soit K un corps, et soit E un K-espace vectoriel non nul. Alors, on a $Z\big(\mathrm{GL}(E)\big) = K^{\times}\mathrm{Id}_E$.

Si E est de dimension finie $n \geqslant 1$, on a de plus

$$Z\big(\mathrm{SL}(E)\big) = Z\big(\mathrm{GL}(E)\big) \cap \mathrm{SL}(E) = \{\lambda \mathrm{Id}_E \mid \lambda \in K, \quad \text{tel que } \lambda^n = 1\}.$$

Les résultats étant clairs si $\dim_K(E) = 1$, supposons que $\dim_K(E) \geqslant 2$ (éventuellement infinie).

Nous allons en fait montrer un résultat plus fort, à savoir qu'un endomorphisme $u \in \mathrm{GL}(E)$ commutant avec toute transvection est une homothétie.

Soit D une droite de E, et soit $a \in D \setminus \{0\}$. On a donc $D = Ka$. Soit H un hyperplan de E contenant D. Un tel hyperplan existe bien. En effet, comme $\dim_K(E) \geqslant 2$, on peut choisir $x_0 \in E$ tel que la famille (a, x_0) soit libre. On considère alors un supplémentaire H' de $\mathrm{Vect}_K(a, x_0)$ dans E, et l'on pose $H = Ka \oplus H'$. On a alors

$$E \;=\; Ka \oplus Kx_0 \oplus H' \;=\; H \oplus Kx_0.$$

Ainsi, H est un hyperplan contenant D.

Soit $v \in \mathscr{L}(E)$ l'unique application linéaire telle que

$$v_{|H} = \mathrm{Id}_H \quad \text{et} \quad v(x_0) = a + x_0.$$

En utilisant la définition, on vérifie que v est une transvection de droite D et d'hyperplan H. Ainsi, $D = \mathrm{Im}(v - \mathrm{Id}_E)$, par la proposition IV-5.2 (2).

Dans ce cas, on a

$$\mathrm{Im}(u\circ v\circ u^{-1}-\mathrm{Id}_E)=\mathrm{Im}\big(u\circ(v-\mathrm{Id}_E)\circ u^{-1}\big)=u\big(\mathrm{Im}(v-\mathrm{Id}_E)\big)=u(D).$$

Comme $u\circ v\circ u^{-1}=v$ par hypothèse, on aussi $\mathrm{Im}(uvu^{-1}-\mathrm{Id}_E)=D$. Ainsi, pour toute droite D de E, $u(D)=D$. Autrement dit, u laisse stable toute droite de E. Un résultat classique nous dit alors que u est une homothétie.

Remontrons-le rapidement : si $u\in\mathscr{L}(E)$ stabilise toute droite de E, alors pour tout $x\in E$, il existe $\lambda_x\in K$ tel que $u(x)=\lambda_x x$. Soient x, $y\in E\setminus\{0\}$ deux vecteurs non nuls. Si x et y sont linéairement indépendants sur K, la relation $u(x+y)=u(x)+u(y)$ implique que

$$(\lambda_{x+y}-\lambda_x)x+(\lambda_{x+y}-\lambda_y)y=0,$$

d'où $\lambda_x=\lambda_{x+y}=\lambda_y$. Si x et y sont liés, par exemple si $y=\mu x$, $\mu\in K^\times$, la relation $u(y)=\mu u(x)$ montre que $\mu\lambda_y=\mu\lambda_x$, d'où $\lambda_y=\lambda_x$.

Finalement, pour tous x, $y\in E$ non nuls, on a $\lambda_x=\lambda_y$. Si l'on note λ cette valeur commune, on en déduit que $u(x)=\lambda x$, pour tout $x\in E$ non nul. Comme $u(0)=0=\lambda\cdot 0$, on a bien $u=\lambda\mathrm{Id}_E$.

Comme un élément de $Z\big(\mathrm{GL}(E)\big)$ commute en particulier avec toute transvection, on obtient $Z\big(\mathrm{GL}(E)\big)\subset K^\times\mathrm{Id}_E$. Puisque toute homothétie commute évidemment avec tout endomorphisme, on a bien $Z\big(\mathrm{GL}(E)\big)=K^\times\mathrm{Id}_E$.

Supposons E de dimension finie $n\geqslant 2$. Comme les transvections sont des éléments de $\mathrm{SL}(E)$, un élément dans le centre de $\mathrm{SL}(E)$ commute avec toute transvection. C'est donc une homothétie par le point précédent. On obtient alors le résultat en remarquant que $\det(\lambda\mathrm{Id}_E)=\lambda^n$.

Le cas du groupe orthogonal n'est pas plus compliqué.

2.13. Exemple. Soit E un espace euclidien de dimension $n\geqslant 1$. Alors,

$$Z\big(\mathrm{O}(E)\big)=\{\pm\mathrm{Id}_E\}.$$

Avant de passer à la démonstration proprement dite, remarquons que, pour tout $x\in E\setminus\{0\}$, on a

$$u\circ\tau_x\circ u^{-1}=\tau_{u(x)}.$$

En effet, on a

$$(u\circ\tau_x\circ u^{-1})\big(u(x)\big)=u\big(\tau_x(x)\big)=u(-x)=-u(x).$$

De plus, puisque u et u^{-1} sont des isométries, pour tout $y\in E$, les vecteurs x et y sont orthogonaux si, et seulement si, $u(x)$ et $u(y)$ le sont. Autrement dit, on a $\{u(x)\}^\perp=u(\{x\}^\perp)$.

Pour tout y' appartenant à $\{u(x)\}^\perp$, il existe donc un vecteur y dans $\{x\}^\perp$ tel que $y' = u(y)$, et l'on a ainsi

$$(u \circ \tau_x \circ u^{-1})(y') = (u \circ \tau_x \circ u^{-1})\big(u(y)\big) = u\big(\tau_x(y)\big) = u(y) = y',$$

d'où l'égalité souhaitée.

Revenons à notre problème. Clairement, Id_E et $-\mathrm{Id}_E$ sont des isométries commutant avec tout élément de $\mathrm{O}(E)$. Montrons la réciproque. Pour cela, on va montrer le fait plus précis suivant : si $u \in \mathrm{O}(E)$ commute avec toute réflexion orthogonale, alors $u = \pm\mathrm{Id}_E$.

Supposons donc que pour tout $x \in E \setminus \{0\}$, on ait $\tau_x \circ u = u \circ \tau_x$. On obtient alors

$$\tau_x = u \circ \tau_x \circ u^{-1} = \tau_{u(x)}, \qquad \text{pour tout } x \in E \setminus \{0\}.$$

Écrivons $u(x) = \lambda_x x + h$, avec $\lambda_x \in \mathbb{R}$ et $h \in \{x\}^\perp$. On a d'une part

$$\tau_{u(x)}\big(u(x)\big) = -u(x) = -\lambda_x x - h,$$

et d'autre part

$$\tau_{u(x)}\big(u(x)\big) = \tau_x\big(u(x)\big) = \tau_x(\lambda_x + h) = -\lambda_x x + h.$$

On en déduit que $h = 0$, soit $u(x) = \lambda_x x$. Ainsi, u stabilise toute droite de E, et u est donc une homothétie d'après l'exemple 2.12. Comme u est une isométrie, on obtient $u = \pm\mathrm{Id}_E$.

Puisqu'un élément de $Z\big(\mathrm{O}(E)\big)$ commute en particulier avec toute réflexion orthogonale, on a fini.

On termine cette liste d'exemples en regardant le cas du groupe spécial orthogonal.

2.14. Exemple. Soit E un espace euclidien de dimension $n \geqslant 1$. Alors, on a

$$Z\big(\mathrm{O}^+(E)\big) = \begin{cases} \mathrm{O}^+(E) & \text{si } n = 1,\ 2 \\ \{\pm\mathrm{Id}_E\} & \text{si } n \geqslant 4 \text{ est pair} \\ \{\mathrm{Id}_E\} & \text{si } n \geqslant 3 \text{ est impair.} \end{cases}$$

Remarquons tout d'abord que Id_E et $-\mathrm{Id}_E$ sont des isométries qui commutent avec tout élément de $\mathrm{O}(E)$. De plus, Id_E est bien entendu une isométrie directe, tandis que $\det(-\mathrm{Id}_E) = (-1)^n$. Ainsi, $-\mathrm{Id}_E \in \mathrm{O}^+(E)$ si, et seulement si, n est pair. Dans ce cas, c'est une isométrie directe qui commute avec tout élément de $\mathrm{O}^+(E)$. Enfin, si $n = 1$ ou 2, le groupe $\mathrm{O}^+(E)$ est abélien, donc égal à son centre.

On peut donc supposer que $n \geqslant 3$ dans la suite. Dans ce cas, on va démontrer que si $u \in \mathrm{O}(E)$ commute avec tout renversement, alors $u = \pm\mathrm{Id}_E$. Le résultat voulu en découlera immédiatement.

Soit P un sous-espace de E de dimension 2, et soit $s \in \mathscr{L}(E)$ l'unique endomorphisme de E tel que

$$s(x) = x, \quad \text{pour tout } x \in P^\perp \quad \text{et} \quad s(x) = -x, \quad \text{pour tout } x \in P.$$

Ainsi, s est une symétrie orthogonale telle que $\mathrm{Ker}(s + \mathrm{Id}_E) = P$. C'est donc un renversement. Par conséquent, on a $s \circ u = u \circ s$ par hypothèse. Autrement dit, $u \circ s \circ u^{-1} = s$. Or, on a

$$\mathrm{Ker}(u \circ s \circ u^{-1} + \mathrm{Id}_E) = \mathrm{Ker}\big(u \circ (s + \mathrm{Id}_E) \circ u^{-1}\big) = u\big(\mathrm{Ker}(s + \mathrm{Id}_E)\big) = u(P),$$

et l'on obtient donc en particulier $u(P) = P$. Ainsi, u stabilise tout plan de E. Montrons alors que u stabilise toute droite de E.

Soit $D = \mathbb{R}x_0$ une droite de E. Puisque $n \geqslant 3$, il existe x_1, $x_2 \in E$ tels que la famille (x_0, x_1, x_2) soit libre. Les sous-espaces

$$P_1 = \mathrm{Vect}_{\mathbb{R}}(x_0, x_1) \quad \text{et} \quad P_2 = \mathrm{Vect}_{\mathbb{R}}(x_0, x_2)$$

sont deux plans de E tels que $P_1 \cap P_2 = D$. Par conséquent

$$u(D) = u(P_1 \cap P_2) = u(P_1) \cap u(P_2) = P_1 \cap P_2 = D,$$

la deuxième égalité découlant de la bijectivité de u. Ainsi, u stabilise toute droite de E, et u est donc une homothétie, d'après l'exemple 2.12. Comme u est une isométrie, on obtient $u = \pm\mathrm{Id}_E$.

Le lemme suivant fournit toute une famille d'exemples de sous-groupes distingués.

2.15. Lemme. *Soit $f : G \longrightarrow G'$ un morphisme de groupes. Alors,* $\mathrm{Ker}(f)$ *est un sous-groupe distingué de G.*

Démonstration. On sait déjà que c'est un sous-groupe de G. Soit $g \in G$ et soit $x \in \mathrm{Ker}(f)$. On a

$$f(gxg^{-1}) = f(g)f(x)f(g)^{-1} = f(g)1_{G'}f(g)^{-1} = 1_{G'},$$

et donc $gxg^{-1} \in \mathrm{Ker}(f)$, ce qu'il fallait vérifier. □

2.16. Remarque. D'après ce lemme, pour démontrer qu'une partie H d'un groupe G est un sous-groupe distingué de G, on peut essayer de l'identifier au noyau d'un morphisme de groupes $f : G \longrightarrow G'$.

2.17. Exemples

(1) Soit E un K-espace vectoriel de dimension finie. L'ensemble $\mathrm{SL}(E)$ des automorphismes de l'espace vectoriel E et de déterminant 1 est un sous-groupe distingué de $\mathrm{GL}(E)$, puisque c'est le noyau du morphisme déterminant $\det : \mathrm{GL}(E) \longrightarrow K^\times$.

(2) Pour tout $n \geqslant 2$, le groupe alterné $\mathfrak{A}_n$ est un sous-groupe distingué du groupe symétrique $\mathfrak{S}_n$, puisque c'est le noyau du morphisme signature $\varepsilon : \mathfrak{S}_n \longrightarrow \{\pm 1\}$.

Une notion centrale en théorie des groupes est celle de groupe simple. Les groupes simples sont, en quelque sorte, les briques élémentaires dont « tout groupe fini est constitué ».

2.18. Définition. On dit qu'un groupe G est *simple* si $G \neq \{1_G\}$, et si G n'a pas de sous-groupe propre qui soit distingué dans G.

Autrement dit, G est simple si $G \neq \{1_G\}$, et si $\{1_G\}$ et G sont les seuls sous-groupes distingués de G.

2.19. Exemples

(1) Si E est un K-espace vectoriel de dimension finie, le groupe $\mathrm{GL}(E)$ n'est pas simple, puisque $\mathrm{SL}(E)$ est un sous-groupe propre distingué dans $\mathrm{GL}(E)$ (sauf cas exceptionnels).

(2) Le groupe $\mathfrak{S}_n$ n'est pas simple pour $n \geqslant 3$, puisque $\mathfrak{A}_n$ est un sous-groupe propre distingué dans $\mathfrak{S}_n$.

(3) En revanche, $\mathfrak{S}_2$ est simple, comme le lecteur le vérifiera immédiatement.

(4) Un groupe non abélien dont le centre est non trivial n'est pas simple. C'est le cas par exemple de $\mathrm{SL}(E)$ lorsque E est un K-espace vectoriel de dimension $n \geqslant 2$, dès que l'équation

$$\lambda^n = 1, \quad \lambda \in K$$

possède au moins une solution $\lambda \neq 1$ dans K (comme par exemple lorsque $K = \mathbb{C}$).

(5) Les groupes abéliens simples sont les groupes $\mathbb{Z}/p\mathbb{Z}$, où p est premier, à isomorphisme près.

Cela sera démontré dans l'exercice 21.

Nous verrons dans des chapitres ultérieurs des familles de groupes simples.

2.20. Remarque. La classification des groupes finis simples à isomorphisme près est connue. Son obtention a été un travail de longue haleine, et la dernière version de la démonstration de cette classification nécessite environ cinq mille pages.

3. Sous-groupes engendrés par une partie

Nous avons vu que la notion de sous-groupe ne se comporte pas bien vis-à-vis de la réunion. En revanche, tout se passe bien par intersection.

3.1. Lemme. *Soit G un groupe, et soit $(H_i)_{i\in I}$ une famille non vide de sous-groupes de G. Alors, $\bigcap_{i\in I} H_i$ est un sous-groupe de G.*

Démonstration. Posons donc $H = \bigcap_{i\in I} H_i$. Puisque $1_G \in H_i$ pour tout $i \in I$, alors $1_G \in H$. En particulier, H est non vide. Si $x, y \in H$, alors $x, y \in H_i$ pour tout $i \in I$. On a alors $xy^{-1} \in H_i$ pour tout $i \in I$ puisque H_i est un sous-groupe de G, et donc $xy^{-1} \in H$. □

3.2. Définition. Soit G un groupe, et soit P une partie de G (qui peut être éventuellement vide). Le *sous-groupe de G engendré par P*, noté $\langle P\rangle$, est l'intersection de tous les sous-groupes de G contenant P.

Autrement dit, $\langle P\rangle$ est le plus petit sous-groupe de G contenant P (au sens de l'inclusion).

Cette définition a bien un sens, puisque d'une part on intersecte sur une famille non vide (elle contient au moins G), et le lemme précédent nous dit que l'on obtient bien un sous-groupe.

3.3. Remarques. Soit G un groupe et soient P, Q deux parties de G. Alors :

(1) on a $P \subset \langle P\rangle$;

(2) soit H un sous-groupe de G. Si $P \subset H$, alors $\langle P\rangle \subset H$;

(3) on a $\langle P\rangle = P$ si, et seulement si, P est un sous-groupe de G ;

(4) si $P \subset Q$, alors $\langle P\rangle \subset \langle Q\rangle$;

(5) on a $\langle\emptyset\rangle = \{1_G\}$.

On a une description précise de $\langle P\rangle$.

3.4. Théorème. *Soit G un groupe, et soit P une partie de G. On note*

$$P^{-1} = \{x^{-1} \mid x \in P\}.$$

Alors, on a

$$\langle P\rangle = \{x_1 \cdots x_n \mid n \geqslant 0,\ x_i \in P \cup P^{-1}\}.$$

Démonstration. Soit $H = \{x_1 \cdots x_n \mid n \geqslant 0,\ x_i \in P \cup P^{-1}\}$. Montrons tout d'abord que $H \subset \langle P\rangle$.

Si $x \in P \subset \langle P \rangle$, alors $x^{-1} \in \langle P \rangle$, car $\langle P \rangle$ est un sous-groupe de G. Par suite, $P \cup P^{-1} \subset \langle P \rangle$, et comme $\langle P \rangle$ est un sous-groupe, on a $x_1 \cdots x_n \in \langle P \rangle$, pour tous $x_1, \ldots, x_n \in P \cup P^{-1}$ (même si $n = 0$, puisque le neutre de G est dans tout sous-groupe). Il en découle que $H \subset \langle P \rangle$.

Pour montrer que $\langle P \rangle \subset H$, il suffit de montrer que H est un sous-groupe de G contenant P, puisque $\langle P \rangle$ est le plus petit sous-groupe de G vérifiant cette propriété. En prenant $n = 1$ et $x_1 \in P$, on voit que $P \subset H$. En prenant cette fois $n = 0$, on obtient $1_G \in H$. En particulier, H est non vide. Par ailleurs, si $x \in P \cup P^{-1}$, on a $x^{-1} \in P \cup P^{-1}$. Par conséquent, si $x = x_1 \cdots x_n$, $y = y_1 \cdots y_m$ sont des éléments de H, on en déduit que

$$xy^{-1} = x_1 \cdots x_n y_m^{-1} \cdots y_1^{-1} \in H.$$

Ainsi, H est un sous-groupe de G contenant P, et l'on a fini. □

Notation. Si $P = \{a_1, \ldots, a_r\}$, le sous-groupe engendré par P est simplement noté $\langle a_1, \ldots, a_r \rangle$.

3.5. Exemple. Soit G un groupe, et soit $x \in G$. Alors, on a

$$\langle x \rangle = \{x^n \mid n \in \mathbb{Z}\}.$$

En effet, le théorème précédent nous dit que

$$\langle x \rangle = \{x^{\varepsilon_1} \cdots x^{\varepsilon_n} \mid n \geqslant 0,\ \varepsilon_i = \pm 1\}.$$

En prenant les ε_i tous égaux à 1 ou -1, on obtient l'inclusion « $\supset$ ». D'autre part, on a

$$x^{\varepsilon_1} \cdots x^{\varepsilon_n} = x^{\varepsilon_1 + \cdots + \varepsilon_n}.$$

Comme $\varepsilon_1 + \cdots + \varepsilon_n \in \mathbb{Z}$, on a l'inclusion manquante.

En particulier, $\langle x \rangle$ est toujours un groupe abélien, puisque

$$x^n x^m = x^{n+m} = x^{m+n} = x^m x^n.$$

3.6. Remarque. Si G est **abélien**, et si $P = \{a_1, \ldots, a_r\}$, la description de $\langle P \rangle$ est un peu plus simple . En effet, on a

$$\langle a_1, \ldots, a_r \rangle = \{a_1^{m_1} \cdots a_r^{m_r} \mid m_1, \ldots, m_r \in \mathbb{Z}\}.$$

En version additive, cela donne

$$\langle a_1, \ldots, a_r \rangle = \{m_1 \cdot a_1 + \cdots + m_r \cdot a_r \mid m_1, \ldots, m_r \in \mathbb{Z}\}.$$

3.7. Définition. Soit G un groupe, et soit P une partie de G. On dit que P *engendre* G si $G = \langle P \rangle$. On dit aussi que P est une partie génératrice, ou un système de générateurs de G. Par exemple, P engendre $\langle P \rangle$.

3.8. Exemples

(1) Si $n \geqslant 3$, le groupe D_n des isométries d'un n-gone régulier est un groupe d'ordre $2n$. Si O est le centre du n-gone, et si S en est un sommet fixé, D_n est engendré par une rotation σ de mesure d'angle $\frac{2\pi}{n}$ et de centre O, et par une réflexion orthogonale d'axe (OS). Ces résultats seront démontrés dans l'exercice 40.

(2) Soit $f : G \longrightarrow G'$ un morphisme de groupes. Pour toute partie P de G, on a $f(\langle P\rangle) = \langle f(P)\rangle$ (exercice laissé au lecteur).

(3) Soit K un corps et soit $n \geqslant 1$. La proposition IV-5.8 montre en particulier que $\mathrm{SL}_n(K)$ est engendré par les matrices de transvections, et que $\mathrm{GL}_n(K)$ est engendré par les matrices de transvections et les matrices de dilatations.

(4) Si K est un corps possédant au moins trois éléments, et si E est un espace vectoriel non nul de dimension finie sur K, les dilatations engendrent $\mathrm{GL}(E)$, d'après la proposition IV-5.9.

(5) Soit E un espace euclidien. Le théorème V-3.7 montre que $\mathrm{O}(E)$ est engendré par les réflexions orthogonales. Si de plus $\dim_{\mathbb{R}}(E) \geqslant 3$, le théorème V-3.9 montre que $\mathrm{O}^+(E)$ est engendré par les renversements.

(6) Le lemme IV-1.2 montre en particulier que pour tout entier $n \geqslant 1$, le groupe $\mathfrak{S}_n$ est engendré par les transpositions.

3.9. Remarque. La notion de partie génératrice est très utile pour simplifier les calculs.

Par exemple, supposons que H et H' soient deux groupes, et que l'on veuille montrer que $H \subset H'$. Si l'on dispose d'une partie génératrice P de H, alors il faut et il suffit de montrer que $P \subset H'$.

De même, supposons que H soit un sous-groupe de G, et que H soit engendré par une partie P. Alors, pour montrer que H est un sous-groupe distingué de G, il faut et il suffit de montrer que pour tout $g \in G$, et pour tout $x \in P$, on a $gxg^{-1} \in H$.

En effet, d'après l'exemple 3.8 (2), l'ensemble gHg^{-1} est un sous-groupe de G engendré par $gPg^{-1} = \{gxg^{-1} \mid x \in P\}$. Si la condition précédente est vérifiée, on a alors $\langle gPg^{-1}\rangle = gHg^{-1} \subset H$, et H est distingué dans G.

Ces « astuces de calcul » seront utilisées dans la suite sans référence complémentaire.

Revenons aux sous-groupes du groupe additif $\mathbb{Z}$. On a vu qu'ils étaient tous de la forme $n\mathbb{Z}$. On devrait alors pouvoir décrire l'intersection de deux ou plusieurs sous-groupes et le sous-groupe engendré par une partie de $\mathbb{Z}$ en fonction d'un seul élément.

3.10. Proposition. *Soient $a, b \in \mathbb{Z}$, et soient $d, m \in \mathbb{Z}$ tels que*

$$\langle a, b\rangle = a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z} \quad \text{et} \quad a\mathbb{Z} \cap b\mathbb{Z} = m\mathbb{Z}.$$

Alors, d est un pgcd de a et b, et m est un ppcm de a et b.

Démonstration. L'égalité $\langle a, b\rangle = a\mathbb{Z} + b\mathbb{Z}$ découle de la remarque 3.6, plus haut. Soit $d \in \mathbb{Z}$ tel que $a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}$. Comme $a \in \langle a, b\rangle = d\mathbb{Z}$, a est un multiple de d. Autrement dit, $d \mid a$. De même, $d \mid b$. Si maintenant $c \in \mathbb{Z}$ vérifie $c \mid a$ et $c \mid b$, alors c divise tout élément de $a\mathbb{Z} + b\mathbb{Z} = d\mathbb{Z}$. En particulier, $c \mid d$. On en déduit que d est un pgcd de a et b.

Montrons la deuxième partie de la proposition. Soit $m \in \mathbb{Z}$ tel que

$$a\mathbb{Z} \cap b\mathbb{Z} = m\mathbb{Z}.$$

On a $m \in m\mathbb{Z} = a\mathbb{Z} \cap b\mathbb{Z}$, donc $m \in a\mathbb{Z}$ et $m \in b\mathbb{Z}$. Par conséquent, $a \mid m$ et $b \mid m$. Si maintenant $c \in \mathbb{Z}$ vérifie $a \mid c$ et $b \mid c$, alors $c \in a\mathbb{Z} \cap b\mathbb{Z} = m\mathbb{Z}$, et donc $m \mid c$. On en déduit que m est un ppcm de a et b. □

3.11. Remarque. Cette proposition et sa démonstration se généralisent à une famille $(a_i)_{i\in I}$ d'entiers. On a donc

$$\langle a_i, i \in I\rangle = \operatorname{pgcd}\big((a_i)_{i\in I}\big)\mathbb{Z}.$$

En particulier, il existe une famille $(u_i)_{i\in I}$ d'entiers presque tous nuls tels que

$$\sum_{i\in I} u_i a_i = \operatorname{pgcd}\big((a_i)_{i\in I}\big).$$

Autrement dit, on a encore l'existence d'une relation de Bézout.

Si l'on pose $\operatorname{ppcm}\big((a_i)_{i\in I}\big) = 0$ lorsque $(a_i)_{i\in I}$ n'est pas bornée, on a l'égalité

$$\bigcap_{i\in I} a_i\mathbb{Z} = \operatorname{ppcm}\big((a_i)_{i\in I}\big)\mathbb{Z}$$

dans tous les cas. En effet, si $m \geqslant 0$ vérifie $\bigcap_{i\in I} a_i\mathbb{Z} = m\mathbb{Z}$, on a $a_i \mid m$ pour tout $i \in I$. En particulier, si $m > 0$, alors on a

$$|a_i| \leqslant m, \quad \text{pour tout } i \in I.$$

Ainsi, si $(a_i)_{i\in I}$ n'est pas bornée, $m = 0$.

4. Théorème de Lagrange, ordre d'un élément

Nous allons maintenant étudier l'ordre des sous-groupes d'un groupe G. On commence par associer deux relations d'équivalence à un sous-groupe quelconque H de G.

4.1. Lemme. *Soit G un groupe, et soit H un sous-groupe. Pour tout $a \in G$, on note*

$$aH = \{ah \mid h \in H\} \quad \text{et} \quad Ha = \{ha \mid h \in H\}.$$

(1) *La relation $\sim_g$ sur G définie par*

$$x \sim_g y \quad si \quad x^{-1}y \in H$$

est une relation d'équivalence. De plus, on a

$$\overline{x} = xH, \quad pour\ tout\ x \in G.$$

(2) *La relation $\sim_d$ sur G définie par*

$$x \sim_d y \quad si \quad yx^{-1} \in H$$

est une relation d'équivalence. De plus, on a

$$\overline{x} = Hx, \quad pour\ tout\ x \in G.$$

Enfin, pour tous x, $y \in G$, on a

$$x \sim_g y \iff x^{-1} \sim_d y^{-1}.$$

Démonstration. On démontre le premier point, la preuve du second point étant similaire. Pour tout $x \in G$, on a $x \sim_g x$ puisque $x^{-1}x = 1_G \in H$. Soient x, y, $z \in G$. Si $x^{-1}y \in H$, alors $(x^{-1}y)^{-1} = y^{-1}x \in H$. Autrement dit, si $x \sim_g y$, alors on a aussi $y \sim_g x$. Enfin, si $x^{-1}y, y^{-1}z \in H$, on a aussi $(x^{-1}y)(y^{-1}z) = x^{-1}z \in H$. Ainsi, si $x \sim_g y$ et $y \sim_g z$, alors $x \sim_g z$. L'égalité $\overline{x} = xH$ est claire.

Enfin, pour tous x, $y \in H$, on a

$$x^{-1}y \in H \iff (x^{-1}y)^{-1} \in H \iff y^{-1}(x^{-1})^{-1} \in H,$$

ce qui démontre la dernière assertion. □

4.2. Définition. Soit G un groupe, et soit H un sous-groupe. Un ensemble de la forme $xH, x \in G$ est appelé une *classe à gauche modulo H*. L'ensemble des classes à gauche modulo H est noté G/H.

Un ensemble de la forme $Hx, x \in G$ est appelé une *classe à droite modulo H*. L'ensemble des classes à droite modulo H est noté $H\backslash G$.

Lorsque G est abélien, classes à gauche et classes à droite coïncident. De plus, si G est noté additivement, la classe à gauche ou à droite de x est l'ensemble

$$x + H = \{x + h \mid h \in H\}.$$

4.3. Remarque. Les classes à gauche (resp. les classes à droite) de G modulo H étant les classes d'équivalence pour une certaine relation d'équivalence sur G, on a immédiatement les propriétés suivantes :

(1) les classes à gauche (resp. à droite) de G modulo H forment une partition de G ;

(2) pour tous x, $y \in G$, on a

$$y \in xH \iff xH = yH \iff x^{-1}y \in H,$$

ainsi que

$$y \in Hx \iff Hx = Hy \iff yx^{-1} \in H.$$

En particulier, pour tout $x \in G$, on a les équivalences

$$xH = H \iff Hx = H \iff x \in H.$$

La dernière partie du lemme 4.1 implique facilement que l'application

$$\begin{aligned} G/H &\longrightarrow H\backslash G \\ xH &\longmapsto Hx^{-1} \end{aligned}$$

est bien définie et bijective.

En particulier, le nombre de classes à gauche modulo H est fini si, et seulement si, le nombre de classes à droite modulo H est fini. Dans ce cas, ces deux nombres sont égaux.

4.4. Définition. Soit G un groupe, et soit H un sous-groupe. Le nombre de classes à gauche modulo H, lorsqu'il est fini, est appelé *l'indice de H dans G*, et est noté $[G : H]$. C'est aussi le nombre de classes à droite modulo H. Si G/H est infini, on pose $[G : H] = +\infty$.

4.5. Remarque. Lorsque G est fini, alors $[G : H]$ est fini pour tout sous-groupe H de G, puisqu'il y a au plus $|G|$ classes d'équivalence distinctes.

4.6. Exemple. Soit $G = \mathbb{Z}$ et soit $H = n\mathbb{Z}$, $n \geqslant 0$. Alors, la relation d'équivalence associée à H n'est rien d'autre que la relation de congruence modulo n, et la classe à gauche de $x \in G$ modulo H est simplement sa classe de congruence modulo n, i.e. $x + n\mathbb{Z} = \{x + kn \mid k \in \mathbb{Z}\}$. L'ensemble quotient est donc [1] $\mathbb{Z}/n\mathbb{Z}$.

Si $n > 0$, on a donc $[\mathbb{Z} : n\mathbb{Z}] = |\mathbb{Z}/n\mathbb{Z}| = n$, par le lemme II-6.5. En revanche, on a $[\mathbb{Z} : 0\mathbb{Z}] = +\infty$, puisque la relation de congruence modulo 0 est simplement la relation d'égalité, et qu'il y a donc une infinité de classes d'équivalence.

1. Le lecteur restera pantois devant la cohérence des notations.

Les notions de classe à gauche et classe à droite permettent de donner une autre caractérisation des sous-groupes distingués d'un groupe G.

4.7. Lemme. *Soit G un groupe, et soit H un sous-groupe. Les propriétés suivantes sont équivalentes :*

(1) *le sous-groupe H est distingué dans G ;*

(2) *pour tout $x \in G$, $xH = Hx$;*

(3) *pour tout $x \in G$, $xHx^{-1} = H$, où $xHx^{-1} = \{xhx^{-1} \mid h \in H\}$.*

Démonstration

(1) $\Longrightarrow$ (2). Supposons que H soit distingué dans G et soit $x \in G$. Pour tout $h \in H$, on a $xhx^{-1} \in H$. Il existe donc $h' \in H$ tel que $xh = h'x$. Ainsi, on obtient $xh \in Hx$, ce qui montre que $xH \subset Hx$.

Mais, on a aussi $x^{-1}hx = x^{-1}h(x^{-1})^{-1} \in H$, et donc il existe $h'' \in H$ tel que $hx = xh'' \in xH$. Ainsi, $Hx \subset xH$, d'où finalement $xH = Hx$.

(2) $\Longrightarrow$ (3). Soit $x \in G$. Alors, $xH = Hx$ par hypothèse. En utilisant les définitions, on obtient facilement les égalités

$$xHx^{-1} = (xH)x^{-1} = (Hx)x^{-1} = H(xx^{-1}) = H1_G = H,$$

d'où le résultat voulu.

(3) $\Longrightarrow$ (1). Évident. □

À cette occasion, citons un classique de théorie des groupes.

4.8. Proposition. *Soit G un groupe. Alors, tout sous-groupe d'indice 2 est distingué dans G.*

Démonstration. Soit H un sous-groupe d'indice 2. D'après le lemme précédent, il suffit de montrer que $xH = Hx$, pour tout $x \in G$. Si $x \in H$, c'est clair, puisque dans ce cas, $xH = H = Hx$, d'après la remarque 4.3.

Supposons maintenant que $x \notin H$. Toujours par cette même remarque, on a $xH \neq H$ et $Hx \neq H$. Ainsi, H et xH sont les deux classes à gauche de G modulo H, tandis que H et Hx sont les deux classes à droite de H modulo H. On a donc

$$G = H \cup Hx = H \cup xH,$$

les deux réunions étant disjointes. Mais alors, on en déduit

$$G \setminus H = Hx = xH,$$

d'où le résultat. □

On passe maintenant au théorème de Lagrange.

4.9. Théorème (Lagrange). *Soit G un groupe* **fini**, *et soit H un sous-groupe de G. Alors, on a*

$$|G| = [G : H]\times|H|,$$

soit encore $|G/H| = \dfrac{|G|}{|H|}$. *En particulier, $|H|$ divise $|G|$.*

Démonstration. Soit $r = [G : H]$. Les classes à gauche $x_1H, \ldots, x_rH$ formant une partition de G, on a donc

$$G = x_1H \cup \cdots \cup x_rH,$$

et la réunion est disjointe. On a donc

$$|G| = |x_1H| + \cdots + |x_rH|.$$

Mais, pour tout $a \in G$, l'application

$$\begin{aligned} H &\longrightarrow aH \\ h &\longmapsto ah \end{aligned}$$

est bijective. En effet, elle est surjective par définition de aH, et injective puisque la loi de groupe est simplifiable. En particulier, on a

$$|aH| = |H|, \quad \text{pour tout } a \in G.$$

Ainsi, on obtient

$$|G| = r|H| = [G : H]\times|H|$$

ainsi que $|H| \mid |G|$. □

4.10. Remarque. On peut démontrer que le groupe $\mathfrak{A}_4$, qui est d'ordre 12, n'a pas de sous-groupe d'ordre 6 (cf. exercice VIII-16). Ainsi, cette réciproque du théorème de Lagrange est fausse.

On continue ce paragraphe par une propriété de multiplicativité de l'indice.

4.11. Proposition. *Soit G un groupe, soit H un sous-groupe de G et soit K un sous-groupe de H.*

Alors, $[G : K]$ est fini si, et seulement si, $[G : H]$ et $[H : K]$ sont finis. Dans ce cas, on a

$$[G : K] = [G : H]\times[H : K].$$

Démonstration. Écrivons

$$G/H = \{g_iH \mid i \in I\},$$

où les classes à gauche g_iH, $i \in I$, sont deux à deux distinctes. De même, écrivons

$$H/K = \{h_jK \mid j \in J\},$$

où les classes à gauche h_jK, $j \in J$, sont deux à deux distinctes.

On va montrer que l'on a

$$G/K = \{g_i h_j K \mid (i,j) \in I \times J\}.$$

Soit $g \in G$. Alors, il existe $i \in I$ tel que $g \in g_i H$. On a donc $g = g_i h$ pour un certain $h \in H$. Mais alors, il existe $j \in J$ tel que $h \in h_j K$. Par conséquent, il existe un indice $k \in K$ tel que $h = h_j k$, et donc $g = g_i h_j k$. On en déduit que $g \sim g_i h_j$, et donc que $gK = g_i h_j K$. Ainsi, les classes

$$g_i h_j K, (i,j) \in I \times J$$

sont les diverses classes à gauche modulo K. Il reste à voir qu'elles sont distinctes.

Supposons que l'on ait $g_i h_j K = g_{i'} h_{j'} K$ pour (i,j), $(i',j') \in I \times J$. Alors, il existe $k \in K$ tel que $g_i h_j = g_{i'} h_{j'} k$. On a donc

$$g_i = g_{i'} h_{j'} k h_j^{-1}.$$

Comme $K \subset H$, on obtient $g_i \in g_{i'} H$, et donc $g_i H = g_{i'} H$. Par choix de I, on obtient $i = i'$. On a alors $h_j = h_{j'} k$, et donc $h_j K = h_{j'} K$. Par choix de J, on obtient $j = j'$. Cela montre que les classes $g_i h_j K$, $(i,j) \in I \times J$ sont toutes distinctes.

En particulier, G/K est fini si, et seulement si, $I \times J$ est fini, c'est-à-dire si, et seulement si, I et J sont finis. Dans ce cas, $[G : H]$ et $[H : K]$ sont donc finis et l'on a

$$[G : K] = |I \times J| = |I| \times |J| = [G : H] \times [H : K].$$

Cela achève la démonstration. □

Nous allons maintenant définir l'ordre d'un élément d'un groupe.

4.12. Définition. Soit G un groupe. On dit que $x \in G$ est *d'ordre fini* si le sous-groupe $\langle x \rangle$ est fini. Dans ce cas, *l'ordre* de x est par définition l'ordre de $\langle x \rangle$. On le note $o(x)$.

En particulier, $o(x)$ est bien défini, pour tout $x \in G$, lorsque G est fini.

Cette définition n'est a priori pas très commode à manipuler. Le résultat suivant nous fournit un moyen pratique de calculer l'ordre d'un élément.

4.13. Théorème. *Soit G un groupe, et soit $x \in G$. Alors, x est d'ordre fini si, et seulement si, il existe $m > 0$ tel que $x^m = 1_G$. Dans ce cas, l'ordre de x est le plus petit entier $n > 0$ vérifiant $x^n = 1_G$, et l'on a*

$$\langle x \rangle = \{1_G, x, \ldots, x^{n-1}\}.$$

De plus, pour tout $m \in \mathbb{Z}$, on a

$$x^m = 1_G \iff o(x) \mid m.$$

Enfin, si G est un groupe fini, on a $o(x) \mid |G|$.

Démonstration. Supposons que x soit d'ordre fini, c'est-à-dire que $\langle x \rangle$ soit fini. Alors, le morphisme de groupes

$$\begin{aligned} h_x \colon \mathbb{Z} &\longrightarrow G \\ m &\longmapsto x^m \end{aligned}$$

ne peut être injectif, car sinon h_x induirait un isomorphisme de $\mathbb{Z}$ sur son image $\langle x \rangle$, et $\langle x \rangle$ serait alors infini. Ainsi, $\mathrm{Ker}(h_x)$ est non nul, et donc contient un élément $m > 0$. On a alors $x^m = 1_G$.

Réciproquement, supposons qu'un tel entier $m > 0$ existe, c'est-à-dire que le noyau de h_x n'est pas réduit à 0. Puisque $\mathrm{Ker}(h_x)$ est un sous-groupe de $\mathbb{Z}$, on a $\mathrm{Ker}(h_x) = n\mathbb{Z}$, où n est le plus petit élément strictement positif de $\mathrm{Ker}(h_x)$, d'après la proposition 2.4. Ainsi pour tout $m \in \mathbb{Z}$, on a

$$x^m = 1_G \iff n \mid m.$$

Montrons alors que $\langle x \rangle = \{1_G, x, \ldots, x^{n-1}\}$. L'inclusion « $\supset$ » est claire. Soit $m \in \mathbb{Z}$, et écrivons $m = qn + r$, $q, r \in \mathbb{Z}$, avec $r \in [\![0, n-1]\!]$. Par conséquent,

$$x^m = x^{qn+r} = (x^n)^q x^r = x^r,$$

ce qui montre l'inclusion « $\subset$ ».

Pour finir, il suffit de montrer que les éléments $1_G, x, \ldots, x^{n-1}$ sont deux à deux distincts. Supposons qu'il existe $i, j \in [\![0, n-1]\!]$ tels que $x^i = x^j$. Alors, on a $x^{i-j} = 1_G$. Autrement dit, $i - j \in \mathrm{Ker}(h_x) = n\mathbb{Z}$, i.e. $i - j$ est un multiple de n. Comme $-n < i - j < n$, on a nécessairement $i - j = 0$, soit $i = j$, d'où le résultat voulu par contraposition.

Le fait que $o(x)$ divise l'ordre du groupe G est une application directe du théorème de Lagrange au sous-groupe $\langle x \rangle$. □

4.14. Remarque. La réciproque du dernier point est fausse.

Par exemple, le groupe $G = \{(1,1), (1,-1), (-1,1), (-1,-1)\} \subset \mathbb{R}^\times \times \mathbb{R}^\times$ est d'ordre 4, mais n'a pas d'élément d'ordre 4.

Un autre exemple est donné par le groupe $\mathfrak{A}_4$. Ce groupe n'ayant pas de sous-groupe d'ordre 6 (cf. remarque 4.10), il n'a a fortiori pas d'élément d'ordre 6, bien que 6 divise l'ordre de $\mathfrak{A}_4$, à savoir 12.

4.15. Corollaire. *Soit G un groupe fini d'ordre $n \geqslant 1$. Pour tout $x \in G$, on a $x^n = 1_G$.*

Démonstration. Par le théorème 4.13, on a $o(x) \mid n$ et par le même théorème, on a alors $x^n = 1_G$. □

4.16. Corollaire. *Soit G un groupe, et soit $x \in G$ un élément d'ordre fini. Alors, pour tout $d \geqslant 1$, l'élément x^d est d'ordre fini, et l'on a*

$$o(x^d) = \frac{o(x)}{\operatorname{pgcd}\big(d, o(x)\big)}.$$

En particulier :

(1) *si $d \mid o(x)$, on a $o(x^d) = \dfrac{o(x)}{d}$;*

(2) *si d est premier avec $o(x)$, on a $o(x^d) = o(x)$.*

Démonstration. Soit $x \in G$ et soit $d \geqslant 1$. Posons $c = \operatorname{pgcd}\big(d, o(x)\big)$, et écrivons $d = rc$, $o(x) = sc$, où r et s sont premiers entre eux. Soit $m \in \mathbb{Z}$. On a

$$(x^d)^m = 1_G \iff x^{dm} = 1_G \iff o(x) \mid dm \iff s \mid rm \iff s \mid m,$$

la dernière équivalence étant une application du lemme de Gauss.

Ainsi, x^d est d'ordre fini, puisqu'en particulier $(x^d)^s = 1_G$, et la plus petite solution > 0 de $(x^d)^m = 1_G$ est s. On a donc

$$o(x^d) = s = \frac{o(x)}{\operatorname{pgcd}\big(d, o(x)\big)}.$$ □

4.17. Proposition. *Soit G un groupe, et soient $x, y \in G$ d'ordre fini tels que $xy = yx$. Alors, xy est d'ordre fini. De plus, on a les propriétés suivantes :*

(1) *supposons que $\langle x \rangle \cap \langle y \rangle = \{1_G\}$. Alors, on a $o(xy) = \operatorname{ppcm}\big(o(x), o(y)\big)$;*

(2) *si $o(x)$ et $o(y)$ sont premiers entre eux, on a $o(xy) = o(x)o(y)$.*

Démonstration. Soit $m = \operatorname{ppcm}\big(o(x), o(y)\big)$. On peut donc écrire

$$m = ro(x) = so(y), \quad r, s \geqslant 1.$$

Puisque $xy = yx$, on a

$$(xy)^m = x^m y^m = (x^{o(x)})^r (y^{o(y)})^s = 1_G.$$

Ainsi, xy est d'ordre fini (divisant m).

Supposons que $\langle x \rangle \cap \langle y \rangle = \{1_G\}$. Il s'agit de montrer que $o(xy) = m$. On sait déjà que $o(xy) \mid m$. Soit $\ell = o(xy)$. On a $(xy)^\ell = x^\ell y^\ell = 1_G$.

Ainsi, $x^\ell = y^{-\ell}$, et donc $x^\ell \in \langle x\rangle \cap \langle y\rangle$. Comme $\langle x\rangle \cap \langle y\rangle = \{1_G\}$, on obtient

$$x^\ell = 1_G = y^{-\ell}.$$

On en déduit $o(x)|\ell$ et $o(y)|\ell$. Ainsi, $m|\ell$ par définition du ppcm. On a donc l'égalité souhaitée.

Montrons le dernier point. Comme $\langle x\rangle \cap \langle y\rangle$ est un sous-groupe de $\langle x\rangle$ et de $\langle y\rangle$, son ordre divise à la fois $o(x)$ et $o(y)$. Puisqu'ils sont premiers entre eux, on en déduit que $\langle x\rangle \cap \langle y\rangle$ est d'ordre 1, et donc réduit à 1_G. Par le point précédent, on obtient

$$o(xy) = \mathrm{ppcm}\big(o(x), o(y)\big) = o(x)\times o(y),$$

la dernière égalité provenant du fait que $o(x)$ et $o(y)$ sont premiers entre eux. □

5. Interlude : formule d'inversion de Möbius

Nous faisons maintenant une pause pour démontrer une formule qui paraît anecdotique au premier abord, mais qui nous sera fort utile à plusieurs reprises.

Notation. Dans la suite, $\mathbb{N}^*$ désignera l'ensemble des entiers naturels non nuls.

On commence par définir la fonction de Möbius.

5.1. Définition. La *fonction de Möbius* est la fonction $\mu : \mathbb{N}^* \longrightarrow \mathbb{Z}$ définie par

$$\mu(n) = \begin{cases} (-1)^r & \text{si } n \text{ est produit de } r \text{ nombres premiers distincts} \\ 0 & \text{s'il existe un nombre premier } p \text{ tel que } p^2 \mid n \end{cases}$$

5.2. Exemple. On a $\mu(1) = 1$, $\mu(2) = \mu(3) = -1$, $\mu(18) = 0$ et $\mu(30) = -1$.

Notation. Dans toute la suite, si $n \geqslant 1$ est un entier, les symboles $\sum_{d|n}$ et $\prod_{d|n}$ désigneront une somme ou un produit sur l'ensemble des diviseurs **positifs** d de n.

5.3. Lemme. *Pour tout $n \geqslant 1$, on a*

$$\sum_{d|n} \mu(d) = \begin{cases} 1 & \text{si } n = 1 \\ 0 & \text{si } n \geqslant 2. \end{cases}$$

Démonstration. Soient $p_1, \ldots, p_r$ les diviseurs premiers deux à deux distincts de l'entier n (avec $r \geqslant 0$). Les diviseurs positifs de n qui vont contri-

buer à un terme non nul sont donc alors ceux de la forme

$$p_I = p_{i_1} \times \cdots \times p_{i_s},$$

où $I = \{i_1, \ldots, i_s\}$ est une partie de $[\![1, r]\!]$. On a alors

$$\sum_{d|n} \mu(d) = \sum_I (-1)^{|I|}.$$

Or, pour tout $s \in [\![0, r]\!]$, il y a $\binom{r}{s}$ parties I de $[\![1, r]\!]$ à s éléments. Ainsi, on obtient

$$\sum_{d|n} \mu(d) = \sum_{s=0}^{r} \binom{r}{s} (-1)^s = (1-1)^r,$$

d'où le résultat (en tenant compte du fait que $(1-1)^0 = 1$). □

On peut maintenant énoncer et démontrer la formule d'inversion de Möbius.

5.4. Théorème (Formule d'inversion de Möbius). *Soit G un groupe abélien, et soient f, $g : \mathbb{N}^* \longrightarrow G$ deux applications. On suppose que l'on a*

$$g(n) = \prod_{d|n} f(d), \qquad \textit{pour tout } n \geqslant 1.$$

Alors, on a l'égalité

$$f(n) = \prod_{d|n} g\Big(\frac{n}{d}\Big)^{\mu(d)}, \qquad \textit{pour tout } n \geqslant 1.$$

Démonstration. On remarque tout d'abord que pour tous d, $d' \geqslant 1$, on a l'équivalence

$$d \mid n \text{ et } d' \mid \frac{n}{d} \iff d' \mid n \text{ et } d \mid \frac{n}{d'}.$$

On a alors

$$\begin{aligned} \prod_{d|n} g\Big(\frac{n}{d}\Big)^{\mu(d)} &= \prod_{d|n} \prod_{d'|\frac{n}{d}} f(d')^{\mu(d)} \\ &= \prod_{d'|n} \prod_{d|\frac{n}{d'}} f(d')^{\mu(d)} \\ &= \prod_{d'|n} f(d')^{\sum_{d|\frac{n}{d'}} \mu(d)}. \end{aligned}$$

Par le lemme précédent, la somme $\sum_{d|\frac{n}{d'}} \mu(d)$ est nulle, sauf pour $\frac{n}{d'} = 1$, soit $d' = n$, d'où finalement

$$\prod_{d|n} g\Big(\frac{n}{d}\Big)^{\mu(d)} = f(n),$$

ce qui est le résultat attendu. □

Nous donnerons une application de la formule d'inversion de Möbius dans le prochain paragraphe.

6. Groupes monogènes, groupes cycliques

On finit ce chapitre en introduisant la notion de groupe monogène et de groupe cyclique.

6.1. Définition. On dit que G est *monogène* s'il peut être engendré par un élément, et *cyclique* s'il est monogène fini.

6.2. Exemples

(1) Le groupe $\mathbb{Z}$ est monogène, non cyclique.

(2) Pour tout entier $n \geqslant 1$, le groupe $\mathbb{Z}/n\mathbb{Z}$ est cyclique d'ordre n.

Le théorème suivant élucide la structure des groupes monogènes.

6.3. Théorème

(1) *Tout groupe monogène infini est isomorphe à* $\mathbb{Z}$.

(2) *Tout groupe cyclique d'ordre* n *est isomorphe à* $\mathbb{Z}/n\mathbb{Z}$. *En particulier, deux groupes cycliques sont isomorphes si, et seulement si, ils ont même ordre.*

Démonstration. Soit $x \in G$ un générateur de G. On a donc $G = \langle x \rangle$. Considérons le morphisme de groupes

$$\begin{aligned} h_x : \mathbb{Z} &\longrightarrow G \\ m &\longmapsto x^m. \end{aligned}$$

Puisque x engendre G, h_x est surjectif.

Si G est infini, alors x n'est pas d'ordre fini. Dans ce cas, on a $\mathrm{Ker}(h_x)=0$. Sinon, $\mathrm{Ker}(h_x)$ contiendrait un élément strictement positif, et par le théorème 4.13, x serait d'ordre fini. Ainsi, h_x est injectif, donc un isomorphisme. Par conséquent, on a $G \simeq \mathbb{Z}$ dans ce cas.

Si G est cyclique d'ordre n, alors $o(x) = n$ par définition, et l'on a $x^n = 1_G$ par le théorème 4.13. Pour tout $m \in \mathbb{Z}$, on a donc

$$h_x(m + kn) = x^{m+kn} = x^m x^{kn} = x^m (x^n)^k = x^m, \qquad \text{pour tout } k \in \mathbb{Z}.$$

Ainsi, h_x est constante sur les classes de congruence modulo n. Par la proposition I-2.11, on en déduit que h_x induit une application [2]

$$\begin{aligned} \overline{h}_x : \mathbb{Z}/n\mathbb{Z} &\longrightarrow G \\ \overline{m} &\longmapsto x^m. \end{aligned}$$

Comme précédemment, cette application est surjective. Comme c'est une application entre deux ensembles finis à n éléments, elle est bijective. En-

2. Bien entendu, on peut aller un peu plus rapidement en utilisant le théorème de factorisation (théorème IX-2.1), qui sera énoncé et démontré ultérieurement.

fin, $\overline{h}_x$ est un morphisme de groupes, puisque pour tous $\overline{m}_1$, $\overline{m}_2 \in \mathbb{Z}/n\mathbb{Z}$, on a

$$\overline{h}_x(\overline{m}_1 + \overline{m}_2) = \overline{h}_x(\overline{m_1 + m_2}) = x^{m_1+m_2} = x^{m_1}x^{m_2} = \overline{h}_x(\overline{m}_1)\overline{h}_x(\overline{m}_2).$$

L'application $\overline{h}_x$ est donc un isomorphisme, et l'on a donc $G \simeq \mathbb{Z}/n\mathbb{Z}$. Cela achève la démonstration. □

6.4. Corollaire. *Soit p un nombre premier, et soit G un groupe d'ordre p. Alors, G est cyclique. En particulier, $G \simeq \mathbb{Z}/p\mathbb{Z}$.*

Démonstration. Puisque G est d'ordre p, il contient un élément $x \neq 1_G$. En particulier, $o(x) > 1$. Mais par le théorème de Lagrange, $o(x)$ divise p, et donc nécessairement $o(x) = p$. On a donc $\langle x \rangle = G$, et G est cyclique. Il est donc isomorphe à $\mathbb{Z}/p\mathbb{Z}$ d'après le théorème précédent. □

Nous nous intéressons maintenant aux sous-groupes d'un groupe cyclique.

6.5. Théorème. *Soit G un groupe cyclique d'ordre n. Pour tout diviseur positif d de n, il existe un unique sous-groupe H_d d'ordre d, et ce sous-groupe est cyclique.*

De plus, si x_0 est un générateur de G, on a les égalités

$$H_d = \left\langle x_0^{\frac{n}{d}} \right\rangle = \{x \in G \mid x^d = 1_G\}.$$

Démonstration. Gardons les notations de l'énoncé, et posons $H_d = \left\langle x_0^{\frac{n}{d}} \right\rangle$. Par le corollaire 4.16 (1), H_d est un sous-groupe d'ordre d. De plus, H_d est cyclique. Posons maintenant

$$H'_d = \{x \in G \mid x^d = 1_G\},$$

et montrons que $H'_d \subset H_d$. Soit $x \in H'_d$. Alors, il existe un entier $m \in \mathbb{Z}$ tel que $x = x_0^m$. Par hypothèse, on a

$$x^d = x_0^{md} = 1_G.$$

Puisque $o(x_0) = n$, on a donc $n \mid md$, soit $\frac{n}{d} \mid m$. Ainsi, $x \in H_d$, ce qui démontre que $H'_d \subset H_d$.

Si maintenant H est un sous-groupe de G d'ordre d, le corollaire 4.15 implique que $H \subset H'_d$. On a ainsi

$$H \subset H'_d \subset H_d,$$

et comme H et H_d ont tous les deux d éléments, ces inclusions sont des égalités. Cela achève la démonstration. □

Nous achevons le présent paragraphe (et ce chapitre) en étudiant ci-après l'indicatrice d'Euler.

6.6. Définition. Pour tout $n \geqslant 1$, on note $\varphi(n)$ le nombre d'entiers de l'ensemble $[\![1, n]\!]$ qui sont premiers à n. La fonction $\varphi : \mathbb{N}^* \longrightarrow \mathbb{C}$ est appelée *fonction indicatrice d'Euler.*

6.7. Corollaire. *Pour tout $n \geqslant 1$, et tout $d \mid n$, $\varphi(d)$ est le nombre d'éléments d'ordre d dans un groupe cyclique d'ordre n.*

De plus, pour tout $n \geqslant 1$, on a

$$\varphi(n) = (p_1^{m_1} - p_1^{m_1-1}) \times \cdots \times (p_r^{m_r} - p_r^{m_r-1}),$$

où $n = p_1^{m_1} \times \cdots \times p_r^{m_r}$ est la décomposition en facteurs premiers de n.

Démonstration. Soit G un groupe cyclique d'ordre n, engendré par $x_0 \in G$. Un élément $x \in G$ d'ordre d appartient en particulier à l'ensemble

$$\{x \in G \mid x^d = 1_G\},$$

qui est égal au sous-groupe $\langle x_0^{\frac{n}{d}} \rangle$ d'après le théorème précédent. Il existe donc un unique entier $k \in [\![1, d]\!]$ tel que $x = x_0^{k\frac{n}{d}}$ par le théorème 4.13 (en tenant compte du fait que $x_0^n = x_0^0 = 1_G$). Par le corollaire 4.16, on obtient alors

$$o(x) = o\big((x_0^{\frac{n}{d}})^k\big) = \frac{d}{\operatorname{pgcd}(k, d)},$$

qui est égal à d si, et seulement si, k et d sont premiers entre eux.

Montrons maintenant le second point. En regroupant les éléments de G selon leur ordre, on obtient l'égalité

$$n = \sum_{d \mid n} \varphi(d),$$

et cette égalité est valable pour tout entier $n \geqslant 1$. Par la formule d'inversion de Möbius, on a alors

$$\varphi(n) = \sum_{d \mid n} \mu(d)\, \frac{n}{d}, \qquad \text{pour tout } n \geqslant 1.$$

Si $n = p_1^{m_1} \cdots p_r^{m_r}$ est la décomposition en facteurs premiers de n, et si I décrit l'ensemble des parties de $[\![1, r]\!]$, on a donc

$$\varphi(n) = n \sum_I \frac{(-1)^{|I|}}{p_I},$$

où $p_I = \prod_{i \in I} p_i$. Or, la somme $\sum_I \frac{(-1)^{|I|}}{p_I}$ n'est rien d'autre que ce que l'on obtient en développant le produit

$$\big(1 - \tfrac{1}{p_1}\big) \times \cdots \times \big(1 - \tfrac{1}{p_r}\big).$$

En effet, pour écrire un terme du développement de ce produit, il faut et il suffit de déterminer dans quels facteurs on choisit un terme « $-\frac{1}{p_i}$ », ce qui revient à choisir une partie I de $[\![1, r]\!]$. On a ainsi

$$\varphi(n) = n\big(1 - \tfrac{1}{p_1}\big) \times \cdots \times \big(1 - \tfrac{1}{p_r}\big) = (p_1^{m_1} - p_1^{m_1 - 1}) \times \cdots \times (p_r^{m_r} - p_1^{m_r - 1}),$$

ce qui achève la démonstration. □

7. Exercices

E1. Détailler les exemples du point 1.2.

E2.

a. Justifier que la table de la loi de multiplication dans un groupe fini est un carré latin, i.e. un tableau carré dans lequel chaque élément apparaît une et une seule fois dans chaque ligne et chaque colonne.

b. Donner la liste de toutes les tables de multiplication possibles pour les groupes d'ordre 4, et en déduire qu'un tel groupe est commutatif. Donner des exemples de groupes pour chaque table possible.

c. Si $E = \{1, \ldots, n\}$, le groupe symétrique $\mathfrak{S}(E)$ est noté $\mathfrak{S}_n$. Lister les éléments de $\mathfrak{S}_3$, et écrire la table de la loi de groupe. Ce groupe est-il commutatif ?

E3. Dans le plan euclidien, on considère le carré de sommets

$$A = (-1, 1), \quad B = (-1, -1), \quad C = (1, -1), \quad D = (1, 1).$$

a. Montrer que l'ensemble des isométries vectorielles qui préservent globalement le carré $ABCD$ est un groupe à huit éléments.

On le note D_4.

Indication. On pourra justifier que l'ensemble des deux diagonales est préservé globalement.

b. Écrire les matrices représentatives des éléments de D_4 dans la base canonique $(\vec{i}, \vec{j})$ de $\mathbb{R}^2$.

c. Décrire les éléments de D_4 comme des permutations de $\{A, B, C, D\}$.

On obtient ainsi trois réalisations concrètes du même groupe.

d. Reprendre les questions précédentes en remplaçant le carré par un triangle équilatéral centré en l'origine.

E4. Soit $n \geqslant 1$. On note $(\mathbb{Z}/n\mathbb{Z})^\times$ l'ensemble des éléments de $\mathbb{Z}/n\mathbb{Z}$ inversibles pour la multiplication. Montrer que $\big((\mathbb{Z}/n\mathbb{Z})^\times, \cdot\big)$ est un groupe.

E5. Soit I un intervalle de $\mathbb{R}$ et soit $f : I \longrightarrow \mathbb{R}$ une bijection. On définit une loi $*$ sur I par

$$x * y = f^{-1}\big(f(x) + f(y)\big), \quad \text{pour tous } x, y \in I.$$

a. Montrer que $(I, *)$ est un groupe commutatif. Expliciter le neutre, et l'inverse de $x \in I$ pour la loi $*$.

b. Donner la loi $*$ dans les cas suivants :

(1) $I = \mathbb{R}_+^\times$, $f(x) = \ln(x)$;

(2) $I = \mathbb{R}$, $f(x) = 1 + x$;

(3) $I = \mathbb{R}$, $f(x) = \lambda x$ ($\lambda \in \mathbb{R}^\times$ fixé) ;

(4) $I =]-1, 1[$, $f(x) = \ln\Big(\dfrac{1+x}{1-x}\Big)$.

E6. Si $n, m \geqslant 0$ sont deux entiers naturels, on définit leur nim-somme, notée $n \oplus m$, comme étant l'unique entier obtenu en sommant modulo 2 les chiffres de leur écriture en binaire, sans retenue.

Autrement dit, si $n = \sum_{j \geqslant 0} x_j 2^j$, $m = \sum_{j \geqslant 0} y_j 2^j$, où $x_j, y_j \in \{0, 1\}$ sont presque tous nuls, alors

$$n \oplus m = \sum_{j \geqslant 0} z_j 2^j,$$

où z_j est l'unique élément de $\{0, 1\}$ tel que $z_j \equiv x_j + y_j \mod 2$.

a. Calculer $5 \oplus 0$, $3 \oplus 3$, $4 \oplus 8$, $31 \oplus 26$.

b. Montrer que $(\mathbb{N}, \oplus)$ est un groupe commutatif. Préciser le neutre, et l'opposé de n.

Indication. Pour l'associativité, on pourra montrer le fait suivant : soient $x, y, z \in \mathbb{Z}$. Soient $e, f, g, h \in \{0, 1\}$ tels que

$$e \equiv x + y \ [2], \quad f \equiv e + z \ [2], \quad g \equiv y + z \ [2], \quad h \equiv x + g \ [2].$$

Alors, $f = h$.

Soient $n_1, \ldots, n_p \in \mathbb{N}$.

c. On suppose que $n_1 \oplus \cdots \oplus n_p = 0$. Montrer que pour tout $k \in [\![1, p]\!]$, et pour tout entier $m_k < n_k$, on a

$$n_1 \oplus \cdots \oplus n_{k-1} \oplus m_k \oplus n_{k+1} \oplus \cdots \oplus n_p \neq 0.$$

d. On suppose que $n_1 \oplus \cdots \oplus n_p \neq 0$. Montrer qu'il existe $k \in [\![1, p]\!]$, et un entier $m_k < n_k$ tel que

$$n_1 \oplus \cdots \oplus n_{k-1} \oplus m_k \oplus n_{k+1} \oplus \cdots \oplus n_p = 0.$$

Indication. Soit 2^d la plus grand puissance de 2 intervenant dans l'écriture binaire de $n_1 \oplus \cdots \oplus n_p$. Justifier qu'il existe un entier $k \in [\![1, p]\!]$ tel que le d-ième chiffre de l'écriture binaire de n_k soit 1.

Montrer alors que l'entier $m_k = (n_1 \oplus \cdots \oplus n_p) \oplus n_k$ convient.

e. On considère le jeu à deux joueurs suivant : on se donne un certain nombre de tas d'allumettes. À tour de rôle, chaque joueur choisit un tas, et prend autant d'allumettes qu'il souhaite dans ce tas. Le joueur qui prend la dernière allumette a gagné.

Montrer qu'un des deux joueurs a une stratégie gagnante, et que cela ne dépend que de la configuration de départ. Quelle est cette stratégie ?

E7. Détailler les exemples du point 2.3.

E8. Les ensembles suivants sont-ils des sous-groupes de $(\mathbb{C}, +)$? de $(\mathbb{C}^\times, \cdot)$?

(i) $\mathbb{R}^\times$;

(ii) $\mathbb{R}_+^\times$;

(iii) $\{z \in \mathbb{C} \mid |z| = 2\}$;

(iv) $\mathbb{U}$;

(v) $\mathbb{U}_n$;

(vi) l'ensemble des nombres décimaux ;

(vii) l'ensemble des nombres décimaux non nuls.

Ici, on a noté $\mathbb{U} = \{z \in \mathbb{C} \mid |z| = 1\}$ et $\mathbb{U}_n = \{z \in \mathbb{C} \mid z^n = 1\}$.

E9.

a. Soit G un groupe. On suppose que l'on a $g^2 = 1_G$ pour tout $g \in G$. Montrer que G est commutatif.

b. Soit G l'ensemble des matrices triangulaires supérieures de $\mathrm{GL}_3(\mathbb{F}_3)$ dont les termes diagonaux sont tous égaux à $\overline{1}$.

Montrer que G est un groupe pour la multiplication matricielle, et que l'on a $M^3 = \mathrm{I}_3$, pour tout $M \in G$. Le groupe G est-il commutatif ?

E10. Soient G_1, G_2 deux groupes. Tout sous-groupe de $G_1 \times G_2$ est-il de la forme $H_1 \times H_2$, où H_i est un sous-groupe de G_i ?

E11. Soit $G = \mathrm{GL}_2(\mathbb{C})$. On suppose que l'on a trois matrices A, B, $C \in G$ vérifiant

$$A^2 = B^2 = C^2 = -\mathrm{I}_2, \quad AB = C, \quad BC = A, \quad CA = B.$$

a. Montrer que $\{\pm\mathrm{I}_2, \pm A, \pm B, \pm C\}$ est un groupe pour la multiplication matricielle, et qu'il possède huit éléments. Écrire sa table de multiplication.

Ce groupe s'appelle le groupe quaternionique, et se note $\mathbb{H}_8$ ou Q_8.

b. Expliciter tous les sous-groupes de Q_8.

c. On cherche trois matrices vérifiant les hypothèses de l'exercice. Montrer qu'il suffit de trouver deux matrices A et B telles que

$$A^2 = B^2 = -\mathrm{I}_2 \quad \text{et} \quad AB = -BA.$$

d. Vérifier que $A = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$ et $B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$ conviennent.

E12. Soit G un groupe. On note $Z(G)$ l'ensemble des éléments de G qui commutent avec tout élément de G.

Montrer que $Z(G)$ est un sous-groupe distingué de G, et qu'il est abélien.

E13. Pour $n \in [\![2, 10]\!]$, lister les éléments de $\mathbb{Z}/n\mathbb{Z}$ et $(\mathbb{Z}/n\mathbb{Z})^\times$, et calculer l'ordre de chaque élément.

E14. Calculer l'ordre de tous les éléments du groupe G, et lister tous ses sous-groupes dans les cas suivants :

$$\mathbb{U}_4, \quad \{\pm 1\} \times \{\pm 1\}, \quad \mathfrak{S}_3, \quad D_4.$$

Dire lesquels sont distingués dans G.

Indication. Pour D_4, il y a dix sous-groupes. Il pourra être utile d'écrire les éléments de D_4 sous la forme $\sigma^i\tau^j$, avec $i \in [\![0, 3]\!]$ et $j = 0, 1$, où σ est la rotation de centre O et d'angle de mesure $\frac{\pi}{2}$, et τ est la symétrie orthogonale d'axe Ox.

E15. Soit G un groupe, et soit H un sous-groupe. On pose

$$N_G(H) = \{g \in G \mid gHg^{-1} = H\}.$$

a. Montrer que $N_G(H)$ est un sous-groupe de G, et que H est distingué dans $N_G(H)$.

b. Montrer que $N_G(H)$ est le plus grand sous-groupe de G dans lequel H est distingué.

E16. Soit G un groupe, et soient H et K deux sous-groupes de G.

a. On suppose que $|H|$ et $|K|$ sont premiers entre eux. Calculer $H \cap K$.

b. On suppose que $H \cap K = \{1_G\}$. Montrer que les éléments

$$hk, \quad h \in H,\ k \in K$$

sont distincts. En déduire que si G est fini, alors $|G| \geqslant |H||K|$.

c. On suppose que H et K sont d'ordre p, où p est un nombre premier. Montrer que, soit $H = K$, soit $H \cap K = \{1_G\}$.

E17. Quels sont les sous-groupes finis de $\mathbb{C}^\times$? les éléments d'ordre fini ?

E18. Montrer que les résultats de la proposition 4.17 sont faux si x et y ne commutent pas.

Indication. On pourra considérer $G = \mathrm{GL}_2(\mathbb{C})$, $M = \begin{pmatrix} 0 & -1 \\ 1 & 1 \end{pmatrix}$ et sa transposée.

E19. Vrai ou faux ? Soit G un groupe commutatif.

(i) Le sous-ensemble des éléments d'ordre fini de G est un sous-groupe de G.

(ii) Le sous-ensemble des éléments d'ordre $n \geqslant 1$ (n fixé) est un sous-groupe de G.

(iii) Le sous-ensemble des éléments d'ordre divisant $n \geqslant 1$ (n fixé) est un sous-groupe de G.

(iv) Le sous-ensemble des éléments d'ordre infini de G est un sous-groupe de G.

(v) Si tous les éléments de G sont d'ordre fini, alors G est fini.

Reprendre ensuite les questions précédentes si G n'est plus supposé commutatif.

E20. Soit G un groupe, et soit $g \in G$ un élément d'ordre fini. On suppose que $o(g) = nm$, où $n, m \geqslant 1$ sont premiers entre eux. Montrer qu'il existe deux éléments $x, y \in G$ **uniques** tels que

(1) $o(x) = n,\ o(y) = m$;

(2) $xy = yx$;

(3) $g = xy$.

E21. Soit G un groupe abélien simple.

a. Montrer que G est monogène.

b. Montrer que $\mathbb{Z}$ n'est pas simple.

c. En déduire que G est cyclique, puis que $G \simeq \mathbb{Z}/p\mathbb{Z}$, avec p premier.

d. Montrer que réciproquement, $\mathbb{Z}/p\mathbb{Z}$ est simple, pour tout p premier.

E22. Le groupe $\mathfrak{S}_3$ est-il cyclique ? Déterminer une partie génératrice minimale pour l'inclusion.

E23.

a. Déterminer le sous-groupe de $(\mathbb{Q}, +)$ engendré par $\frac{1}{2}$ et $\frac{1}{3}$.

b. Le groupe $(\mathbb{Q}, +)$ possède-t-il une partie génératrice finie ? (Procéder par l'absurde.)

E24. Quel est le sous-groupe de $\mathrm{O}(\mathbb{R}^2)$ engendré par une rotation ? une réflexion orthogonale ? deux réflexions orthogonales ? une réflexion et une rotation ?

E25. On considère le groupe additif $\mathbb{Z}^2$.

a. Le groupe $\mathbb{Z}^2$ est-il monogène ?

b. Dans quels cas la partie $\{u, v\}$ engendre-t-elle $\mathbb{Z}^2$?

(1) $u = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $v = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$

(2) $u = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$, $v = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

(3) $u = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$, $v = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

c. Soit $\Delta = \{(x, y) \in \mathbb{Z}^2 \mid x + y \in 3\mathbb{Z}, y \in 2\mathbb{Z}\}$. Représenter Δ, montrer que c'est un sous-groupe de $\mathbb{Z}^2$, et donner une partie génératrice minimale pour l'inclusion.

E26. Quel est le sous-groupe de $\mathbb{Z}/24\mathbb{Z}$ engendré par $\overline{6}$ et $\overline{15}$?

E27.

a. Montrer que $(\mathbb{Z}/17\mathbb{Z})^\times$ est un groupe cyclique à seize éléments, et en donner un générateur.

b. Pour $n \leqslant 10$, déterminer si $(\mathbb{Z}/n\mathbb{Z})^\times$ est cyclique ou non, et donner une partie génératrice minimale.

E28.

a. Le groupe $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$ est-il cyclique ?

b. Le groupe $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}$ est-il cyclique ?

Soient $m, n \geqslant 1$ deux entiers, et soit $G = \mathbb{Z}/m\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z}$.

c. Montrer que l'ordre d'un élément de G divise $e = \mathrm{ppcm}(m, n)$. Montrer que $g = (\overline{1}, \overline{1})$ est d'ordre e.

d. Montrer que G est cyclique si, et seulement si, m et n sont premiers entre eux.

E29. Soit G un groupe fini, et soit H un sous-groupe de G. On pose

$$U = \bigcup_{g \in G} gHg^{-1}.$$

a. Soit $r = [G : H]$. Si $g_1H, \ldots, g_rH$ sont les éléments de G/H, montrer que l'on a

$$U \setminus \{1_G\} = \bigcup_{i=1}^{r} (g_iHg_i^{-1} \setminus \{1_G\}).$$

b. En déduire que $|U| \leqslant |G| - [G : H] + 1$.

c. Soit S un sous-ensemble de G contenant au moins un élément de chaque classe de conjugaison dans G. Montrer que S engendre G.

E30. Détailler les exemples du point 1.7.

E31. Les applications suivantes sont-elles des morphismes de groupes ?

(i) $\begin{array}{l} \mathbb{R} \longrightarrow \mathbb{R} \\ x \longmapsto 2x \end{array}$

(ii) $\begin{array}{l} \mathbb{R}^\times \longrightarrow \mathbb{R}^\times \\ x \longmapsto 2x \end{array}$

(iii) $\begin{array}{l} \mathbb{C}^\times \longrightarrow \mathbb{C}^\times \\ x \longmapsto x^2 \end{array}$

(iv) $\begin{array}{l} G \longrightarrow G \\ x \longmapsto x^2 \end{array}$, où G est un groupe

(v) $\begin{array}{l} \mathbb{R} \longrightarrow \mathrm{GL}_2(\mathbb{R}) \\ t \longmapsto \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \end{array}$

(vi) $\begin{array}{l} G \longrightarrow G \\ x \longmapsto x^{-1} \end{array}$, où G est un groupe

E32. Soit $f : G \longrightarrow G'$ un morphisme de groupes.

a. Soit $x \in G$ d'ordre fini. Montrer que $f(x) \in G'$ est d'ordre fini et que l'on a $o\big(f(x)\big) \mid o(x)$.
Si G et G' sont finis, en déduire que

$$o\big(f(x)\big) \mid \mathrm{pgcd}(|G|, |G'|), \qquad \text{pour tout } x \in G.$$

b. On suppose que G et G' sont deux groupes finis d'ordres premiers entre eux. Décrire tous les morphismes de groupes $G \longrightarrow G'$.

c. On suppose que $f : G \longrightarrow G'$ est un isomorphisme de groupes. Comparer $o(x)$ et $o\big(f(x)\big)$. Et si f est supposée seulement injective ?

E33. Soit $f : G \longrightarrow G'$ un morphisme de groupes. Soit $\mathcal{P}$ une partie quelconque de G.

a. Montrer que $f(\langle \mathcal{P} \rangle) = \langle f(\mathcal{P}) \rangle$.

b. Montrer que si $\mathcal{P}$ engendre G, $f(\mathcal{P})$ engendre $\mathrm{Im}(f)$.

c. On suppose que f est un isomorphisme de groupes. Montrer que $\mathcal{P}$ engendre G si, et seulement si, $f(\mathcal{P})$ engendre G'.

E34. Soit G un groupe fini. Quels sont les morphismes de groupes $G \longrightarrow \mathbb{Z}$?

E35. Soient G et G' deux groupes abéliens. On suppose que la loi de G' est notée multiplicativement.

Montrer que l'ensemble $\mathrm{Hom}(G, G')$ des morphismes de groupes de G dans G' est un groupe pour la loi

$$(f_1 \cdot f_2)(g) = f_1(g) \cdot f_2(g), \qquad \text{pour tout } g \in G.$$

E36. Soit G un groupe. On rappelle que, pour tout $g \in G$, l'application

$$\begin{aligned} h_g : \mathbb{Z} &\longrightarrow G \\ m &\longmapsto g^m \end{aligned}$$

est un morphisme de groupes.

a. Montrer que l'application

$$\begin{aligned} \varphi : G &\longrightarrow \mathrm{Hom}(\mathbb{Z}, G) \\ g &\longmapsto h_g \end{aligned}$$

est une bijection, i.e. que tout morphisme de groupes $\mathbb{Z} \longrightarrow G$ est de la forme h_g pour un unique $g \in G$.

b. Si G est commutatif, montrer que φ est un isomorphisme de groupes.

E37. Soit G un groupe cyclique d'ordre n.

a. Soit $m \in \mathbb{Z}$. Montrer que l'application

$$\begin{aligned} \rho_m : G &\longrightarrow G \\ x &\longmapsto x^m \end{aligned}$$

est un morphisme de groupes. Montrer que c'est un automorphisme si, et seulement si, m est premier à n.

b. Montrer que ρ_m est d'ordre d si, et seulement si, $\overline{m} \in (\mathbb{Z}/n\mathbb{Z})^\times$ est d'ordre d.

c. Montrer que tout morphisme $\rho : G \longrightarrow G$ est de la forme ρ_m.

d. Montrer que $\mathrm{Aut}(G) \simeq (\mathbb{Z}/n\mathbb{Z})^\times$.

E38. Soit p un nombre premier. Montrer que pour tout $n \geqslant 1$, on a un isomorphisme de groupes

$$\mathrm{GL}_n(\mathbb{F}_p) \simeq \mathrm{Aut}\big((\mathbb{Z}/p\mathbb{Z})^n\big),$$

qui à une matrice inversible M fait correspondre le morphisme de multiplication à gauche par M.

E39. Soit E un espace euclidien.

a. Donner un morphisme de groupes $\mathrm{O}(E) \longrightarrow \mathbb{C}^\times$ non trivial.

b. Soit $\varphi : \mathrm{O}(E) \longrightarrow \mathbb{C}^\times$ un morphisme de groupes non trivial. Montrer que $\varphi(\tau) = -1$, pour toute réflexion orthogonale τ. En déduire qu'il n'existe qu'un morphisme de groupes $\mathrm{O}(E) \longrightarrow \mathbb{C}^\times$ non trivial.

Indication. Utiliser le fait que les réflexions orthogonales sont conjuguées dans $\mathrm{O}(E)$.

E40. Soit $n \geqslant 3$ un entier. Soit $\mathcal{P}_n$ le n-gone de $\mathbb{R}^2$ de sommets $A_1, \ldots, A_n$, où $A_k = \big(\cos\big(\frac{2k\pi}{n}\big), \sin\big(\frac{2k\pi}{n}\big)\big)$.

a. Montrer que l'ensemble des isométries vectorielles de $\mathbb{R}^2$ laissant globalement invariant $\mathcal{P}_n$ est un groupe d'ordre $2n$, et lister ses éléments. On le note D_n, et on l'appelle le groupe diédral (d'ordre $2n$).

Indication. Une application linéaire est entièrement déterminée par l'image de $\overrightarrow{OA_1}$ et $\overrightarrow{OA_2}$ (pourquoi ?). De plus, les distances et les angles non orientés sont conservés par une isométrie vectorielle.

Soit r la rotation de centre O et d'angle de mesure $\frac{2\pi}{n}$ et soit s la symétrie orthogonale d'axe Ox.

b. Vérifier que r est d'ordre n, que s est d'ordre 2, et que $srs^{-1} = r^{-1}$. Donner les matrices représentatives de r et s dans la base canonique.

c. Montrer que tout élément de D_n s'écrit de manière unique sous la forme $r^i s^j, i \in [\![0, n-1]\!], j = 0, 1$. En déduire que D_n est engendré par r et s.

d. Soit $E = \{A_1, \ldots, A_n\}$ l'ensemble des sommets de $\mathcal{P}_n$. Remarquer qu'un élément de D_n permute les éléments de E, et montrer que l'application

$$\begin{aligned} \varphi : D_n &\longrightarrow \mathfrak{S}(E) \\ f &\longmapsto f_E \end{aligned}$$

est un morphisme de groupes, qui est injectif.

En déduire que D_n est isomorphe à un sous-groupe de $\mathfrak{S}_n$.

e. Déterminer deux éléments $\sigma, \tau \in \mathfrak{S}_n$ tels que σ et τ engendrent un sous-groupe de $\mathfrak{S}_n$ isomorphe à D_n (attention à la parité de n).

Indication. Utiliser l'exercice 33.

f. Le groupe D_n est-il cyclique ?

E41. Soit $n \geqslant 3$. Soit G un groupe. On suppose que G est engendré par deux éléments σ et τ vérifiant

$$o(\sigma) = n, \quad o(\tau) = 2, \quad \tau\sigma\tau^{-1} = \sigma^{-1}.$$

a. Montrer que tout élément de G s'écrit de manière unique sous la forme

$$\sigma^i\tau^j, \quad i \in [\![0, n-1]\!], \ j = 0, 1.$$

Quel est l'ordre de G ?

b. Montrer que G est isomorphe à D_n.

Indication. Utiliser la question précédente et l'exercice 40 pour définir une application $f : G \longrightarrow D_n$ et montrer que c'est un isomorphisme de groupes.

E42. Parmi les groupes suivants, lesquels sont isomorphes ?

$$(\mathbb{R}, +),\ (\mathbb{R}^\times, \cdot),\ (\mathbb{R}_+^\times, \cdot),\ (\mathbb{C}^\times, \cdot),\ \mathbb{U}_6,\ D_3,\ \mathfrak{S}_3,\ \mathbb{Z}/6\mathbb{Z},\ \mathbb{U}_2 \times \mathbb{U}_6.$$

Indication. Regarder les ordres des éléments dans chaque groupe.

E43. Soit G un groupe (noté multiplicativement) et soit p un nombre premier. On suppose que pour tout $g \in G$, on a $g^p = 1_G$.

a. On suppose que G est commutatif.

(1) Montrer que l'application

$$\begin{aligned} \mathbb{Z}/p\mathbb{Z} \times G &\longrightarrow G \\ (\overline{m}, g) &\longmapsto g^m \end{aligned}$$

est bien définie, et munit G d'une structure de $\mathbb{Z}/p\mathbb{Z}$-espace vectoriel.

(2) En déduire que si G est fini, $G \simeq (\mathbb{Z}/p\mathbb{Z})^n$ pour un certain $n \geqslant 0$. En particulier, $|G|$ est une puissance de p.

b. On suppose G quelconque, non trivial.

(1) Justifier l'existence d'un élément $a \in G$ d'ordre p.

(2) On suppose que G n'est pas cyclique d'ordre p. Justifier l'existence de $b \in G$ d'ordre p tel que $b \notin \langle a \rangle$.
Montrer que $\langle a \rangle \cap \langle b \rangle = \{1_G\}$.

c. En déduire que si G est fini et non trivial, on a $|G| = p$ ou $|G| \geqslant p^2$.

Indication. Utiliser l'exercice 16.

E44. Soit G un groupe d'ordre 6.

a. Utiliser l'exercice précédent pour montrer que G contient au moins un élément σ d'ordre 3 et un élément τ d'ordre 2.

Indication. Traiter le cas où G est cyclique à part. Si G n'est pas cyclique, procéder par l'absurde en supposant que G ne contient que des éléments d'ordre 1 ou p ($p = 2, 3$).

b. Montrer que $G = \{1_G,\ \sigma,\ \sigma^2,\ \tau,\ \sigma\tau,\ \sigma^2\tau\}$.

Indication. On pourra utiliser l'exercice 16, ou procéder directement.

c. On suppose que G est commutatif. Montrer que G est isomorphe au groupe $\mathbb{Z}/6\mathbb{Z}$.

d. On suppose que G n'est pas commutatif. Montrer que $\tau\sigma\tau^{-1} = \sigma^{-1}$. En déduire que G est isomorphe à D_3.

e. Résumer le résultat obtenu.

E45. Soit G un groupe d'ordre 8.

a. Montrer qu'un élément $g \in G$ différent du neutre est d'ordre 2, 4 ou 8.

b. On suppose que G possède un élément d'ordre 8. Montrer que G est isomorphe à $\mathbb{Z}/8\mathbb{Z}$.

c. On suppose que G ne possède que des éléments d'ordre 1 ou 2. Utiliser l'exercice 43 pour montrer que G est isomorphe à $(\mathbb{Z}/2\mathbb{Z})^3$.

On suppose maintenant que G n'a pas d'élément d'ordre 8, mais possède au moins un élément $b \in G$ d'ordre 4. On note $H = \langle b \rangle$.

d. Combien y a-t-il de classes à gauche de G modulo H ? En déduire l'existence de $a \in G \setminus H$ tel que $G = H \cup aH$ (la réunion étant disjointe), puis l'égalité

$$G = \{1_G,\ b,\ b^2,\ b^3,\ a,\ ab,\ ab^2,\ ab^3\}.$$

e. (1) Montrer par l'absurde que $a^2 \in H$.

(2) Montrer que si $a^2 = b$ ou b^3, alors a est d'ordre 8.

(3) En déduire que $a^2 = 1_G$ ou $a^2 = b^2$.

f. (1) Montrer par l'absurde que $aba^{-1} \in H$. Quel est l'ordre de aba^{-1} ?

(2) En déduire que $aba^{-1} = b$ ou b^3.

g. On suppose que $aba^{-1} = b$.

(1) Montrer que G est commutatif.

(2) On suppose que $a^2 = 1_G$. Montrer que l'application

$$\begin{aligned} \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z} &\longrightarrow G \\ (\overline{m}, \tilde{n}) &\longmapsto a^m b^n \end{aligned}$$

est bien définie, et est un isomorphisme.

(3) On suppose que $a^2 = b^2$. Montrer que $(ab)^2 = 1_G$. En déduire que l'on a encore $G \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z}$.

h. On suppose que $aba^{-1} = b^3$.

(1) On suppose que $a^2 = 1_G$. En utilisant l'exercice 41, montrer que l'on a $G \simeq D_4$.

(2) On suppose que $a^2 = b^2$. On pose $\omega = b^2$ et $c = ab$. Montrer que l'on a $\omega \in Z(G)$, ainsi que

$$G = \{1_G,\ \omega,\ a,\ \omega a,\ b,\ \omega b,\ c,\ \omega c\}.$$

Vérifier que $a^2 = b^2 = c^2 = \omega$ et que $ab = c,\ bc = a,\ ca = b$.

(3) Montrer que $G \simeq Q_8$, où Q_8 est défini dans l'exercice 11.

i. Montrer que les groupes

$$\mathbb{Z}/8\mathbb{Z},\ \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z},\ (\mathbb{Z}/2\mathbb{Z})^3,\ D_4 \text{ et } Q_8$$

sont deux à deux non isomorphes.

j. Résumer le résultat obtenu.

Chapitre VII

Groupes opérant sur un ensemble

1. Actions de groupe

Nous allons maintenant définir une notion qui est le cœur même de la théorie des groupes, et qui établit un lien naturel entre théorie des groupes et géométrie.

1.1. Définition. Soit E un ensemble et soit G un groupe. Une *action* (ou une *opération*) *à gauche* de G sur E est une application

$$\begin{aligned} G \times E &\longrightarrow E \\ (g, x) &\longmapsto g \cdot x \end{aligned}$$

vérifiant les propriétés suivantes :

(1) pour tout $x \in E$, on a $1_G \cdot x = x$;

(2) pour tous g, $g' \in G$, et tout $x \in E$, on a $g \cdot (g' \cdot x) = gg' \cdot x$.

Dans ce cas, on dit que G *opère* sur E ou *agit* sur E.

1.2. Remarque. Supposons que G agisse (à gauche) sur E. Alors, pour tout $g \in G$, et tout $x \in E$, on a

$$g \cdot (g^{-1} \cdot x) = g^{-1} \cdot (g \cdot x) = x.$$

En particulier, l'application

$$\begin{aligned} \sigma_g : E &\longrightarrow E \\ x &\longmapsto g \cdot x \end{aligned}$$

est bijective, d'inverse

$$\begin{aligned} \sigma_{g^{-1}} : E &\longrightarrow E \\ x &\longmapsto g^{-1} \cdot x. \end{aligned}$$

Autrement dit, on a $\sigma_g \in \mathfrak{S}(E)$. L'application

$$\begin{aligned} \varphi : G &\longrightarrow \mathfrak{S}(E) \\ g &\longmapsto \sigma_g \end{aligned}$$

est alors un morphisme de groupes, puisque pour tous $g, g' \in G$, et pour tout $x \in E$, on a

$$\varphi(gg')(x) = gg'\cdot x = g\cdot(g'\cdot x) = g\cdot\big(\varphi(g')(x)\big) = \big(\varphi(g)\circ\varphi(g')\big)(x).$$

Inversement, si $\psi : G \longrightarrow \mathfrak{S}(E)$ est un morphisme de groupes, on vérifie facilement que

$$\begin{aligned} G \times E &\longrightarrow E \\ (g, x) &\longmapsto g\cdot x = \psi(g)(x) \end{aligned}$$

est une action de G sur E.

On montre aisément que les deux constructions sont inverses l'une de l'autre. Autrement dit, il y a une correspondance bijective entre les actions de G sur E et les morphismes de groupes de G dans $\mathfrak{S}(E)$.

Dans la suite, l'ensemble E sera implicitement supposé non vide.

1.3. Définition. Soit G un groupe opérant sur E. On dit que G agit *fidèlement* sur E si pour tout $g \in G$, on a

$$g\cdot x = x, \quad \text{pour tout } x \in E \quad \Longrightarrow \quad g = 1_G.$$

Autrement dit, G agit fidèlement sur E si le morphisme $\varphi : G \longrightarrow \mathfrak{S}(E)$ associé est injectif.

On dit que le groupe G agit *transitivement* sur E si pour tous x, $x' \in E$, il existe $g \in G$ tel que $x' = g\cdot x$.

1.4. Remarque. Soit G un groupe, et soit H un sous-groupe de G. Toute action d'un groupe G sur un ensemble E induit par restriction une action de H sur E.

1.5. Exemples

(1) Un groupe G agit sur lui-même par translation à gauche, et cette action est fidèle. En effet, pour g, $x \in G$, posons

$$g\cdot x = gx.$$

Alors, on a $1_G\cdot x = 1_G x = x$, et pour tous g, g', $x \in G$, on a

$$g\cdot(g'\cdot x) = g\cdot(g'x) = g(g'x) = (gg')x = gg'\cdot x.$$

Si de plus on a $g\cdot x = gx = x$ pour tout $x \in G$, alors on a $g = 1_G$ en prenant $x = 1_G$.

Cette action de G est également transitive. En effet, pour tous x, y dans G, on a $g \cdot x = y$, avec $g = yx^{-1}$.

(2) Un groupe G agit sur un sous-groupe H distingué dans G, par conjugaison. Autrement dit, l'application

$$\begin{aligned} G \times H &\longrightarrow H \\ (g, x) &\longmapsto g \cdot x = gxg^{-1} \end{aligned}$$

est une action de G sur H.

En particulier, G agit sur lui-même par conjugaison.

Cette action n'est pas fidèle en général. Par exemple, si G est abélien, alors $g \cdot x = gxg^{-1} = x$ pour tout $g \in G$, et tout $x \in H$.

(3) Pour tout ensemble E, le groupe $\mathfrak{S}(E)$ agit sur E par

$$\begin{aligned} \mathfrak{S}(E) \times E &\longrightarrow E \\ (\sigma, x) &\longmapsto \sigma \cdot x = \sigma(x). \end{aligned}$$

Cette action est fidèle et transitive.

(4) Soit $G = \mathrm{GL}_n(\mathbb{C})$ et soit $E = \mathbb{C}^n \setminus \{0\}$. Alors, G agit sur E par

$$\begin{aligned} G \times E &\longrightarrow E \\ (M, v) &\longmapsto Mv, \end{aligned}$$

et cette action est fidèle et transitive.

(5) Soit H un sous-groupe quelconque de G, et soit $E = G/H$ l'ensemble des classes à gauche modulo H. Alors, l'application

$$\begin{aligned} G \times E &\longrightarrow E \\ (g, \overline{g'}) &\longmapsto g \cdot \overline{g'} = \overline{gg'} \end{aligned}$$

est bien définie, et donne une action transitive de G sur E.

Vérifions que l'application est bien définie. Si $g', g'' \in G$ sont deux éléments de G tels que $\overline{g''} = \overline{g'}$, alors il existe $h \in H$ tel que $g'' = g'h$, et l'on a $gg'' = gg'h$. Ainsi, $\overline{gg''} = \overline{gg'}$, ce qu'il fallait montrer. Le fait que l'on obtienne une action transitive de G sur E est alors clair.

(6) Soit G un groupe agissant sur E, et soit $g \in G$. Alors, le groupe $\mathbb{Z}$ agit sur E par

$$\begin{aligned} \mathbb{Z} \times E &\longrightarrow E \\ (m, x) &\longmapsto m * x = g^m \cdot x. \end{aligned}$$

Cette action n'est en général ni fidèle, ni transitive.

(7) Soit $n \geqslant 1$ un entier et soit E un ensemble. Alors, $\mathfrak{S}_n$ agit sur E^n par

$$\begin{aligned} \mathfrak{S}_n \times E^n &\longrightarrow E^n \\ \big(\sigma, (x_1, \ldots, x_n)\big) &\longmapsto (x_{\sigma^{-1}(1)}, \ldots, x_{\sigma^{-1}(n)}). \end{aligned}$$

Le lecteur consciencieux vérifiera que prendre l'inverse σ^{-1} est bien nécessaire pour avoir une action de groupe.

Nous laissons le soin au lecteur de détailler ces exemples, à titre d'exercice.

L'intérêt de faire agir un groupe G sur un ensemble est multiple. Cela peut permettre de comprendre la structure de G en étudiant le morphisme de groupes $\varphi : G \longrightarrow \mathfrak{S}(E)$ associé. Par exemple, le noyau de φ est le sous-groupe des éléments $g \in G$ qui agissent trivialement sur E. Par ailleurs, beaucoup de groupes sont de nature géométrique, et définis comme groupes de transformations préservant globalement une figure géométrique (un bon exemple est donné par le Rubik's cube ou autres puzzles apparentés). Les actions de groupes permettent alors de mieux visualiser certains sous-groupes en les interprétant comme le stabilisateur d'un élément de E.

1.6. Définition. L'ensemble

$$\mathrm{Stab}_G(x) = \{g \in G \mid g{\cdot}x = x\}$$

est appelé le *stabilisateur* de x (sous l'action de G). On vérifie facilement que c'est un sous-groupe de G.

Un autre intérêt majeur d'avoir une action de groupe est qu'elle permet de définir naturellement une relation d'équivalence sur E, comme on va le voir maintenant.

1.7. Lemme. *Soit G opérant sur un ensemble E. La relation sur E définie par*

$$x \sim y \quad \text{s'il existe } g \in G \text{ tel que } y = g{\cdot}x$$

est une relation d'équivalence.

Démonstration. Pour tout $x \in E$, on a $x = 1_G{\cdot}x$, et donc $x \sim x$. Si maintenant $y = g{\cdot}x$ pour un certain $g \in G$, alors

$$g^{-1}{\cdot}y = g^{-1}{\cdot}(g{\cdot}x) = x.$$

Ainsi, on a $y \sim x$ dès que $x \sim y$. Enfin, si $y = g{\cdot}x$ et $z = g'{\cdot}y$, alors

$$z = g'{\cdot}(g{\cdot}x) = g'g{\cdot}x,$$

et la relation est donc transitive. □

1.8. Définition. Soit G opérant sur un ensemble E, et soit $x \in E$. L'ensemble

$$\mathcal{O}_x = \{g \cdot x \mid g \in G\} \subset E$$

est appelé *l'orbite* de x sous l'action de G, ou *G-orbite* de x. C'est la classe d'équivalence de x pour la relation d'équivalence sur E induite par l'action de G sur E.

On dit que $x \in E$ est un *point fixe* sous l'action de G si l'on a $g \cdot x = x$, pour tout $g \in G$. L'ensemble des points fixes est noté E^G.

1.9. Exemple. Si G agit sur lui-même par conjugaison, les orbites correspondantes sont les classes de conjugaison de G.

1.10. Remarques

(1) Soit G un groupe agissant sur E, et soit $x \in E$. Alors, les propriétés suivantes sont équivalentes :
   (i) x est un point fixe sous l'action de G ;
   (ii) $\mathcal{O}_x = \{x\}$;
   (iii) $|\mathcal{O}_x| = 1$;
   (iv) $\mathrm{Stab}_G(x) = G$.

(2) Soit G un groupe agissant sur E. Alors, les propriétés suivantes sont équivalentes :
   (i) G agit transitivement sur E ;
   (ii) pour tout $x \in E$, $\mathcal{O}_x = E$;
   (iii) il existe un $x_0 \in E$ tel que $\mathcal{O}_{x_0} = E$;
   (iv) E n'admet qu'une seule G-orbite, à savoir E lui-même.

Cela est laissé en exercice au lecteur.

Le lemme suivant découle immédiatement de la définition d'une orbite et des propriétés des relations d'équivalence. Il sera constamment utilisé dans la suite, sans référence ultérieure.

1.11. Lemme (Équation aux classes)

Soit G un groupe agissant sur E. Alors, les différentes orbites de E sous l'action de G forment une partition de E. Si de plus E est fini, le cardinal de E est la somme des cardinaux des différentes orbites. Autrement dit, si Ω est l'ensemble des orbites de E sous l'action de G, on a

$$|E| = \sum_{\omega \in \Omega} |\omega|.$$

Nous allons maintenant faire le lien entre orbites et stabilisateurs.

1.12. Proposition. *Soit G un groupe opérant sur E. Pour tout $x \in E$, l'application*

$$\begin{aligned} f : G &\longrightarrow \mathcal{O}_x \\ g &\longmapsto g{\cdot}x \end{aligned}$$

induit par passage au quotient une bijection

$$\begin{aligned} \overline{f} : G/\mathrm{Stab}_G(x) &\xrightarrow{\sim} \mathcal{O}_x \\ \overline{g} &\longmapsto g{\cdot}x. \end{aligned}$$

En particulier, $\mathcal{O}_x$ est un ensemble fini si, et seulement si, $[G : \mathrm{Stab}_G(x)]$ est fini.

Enfin, si G est fini , pour tout $x \in G$, l'ensemble $\mathcal{O}_x$ est fini et son cardinal divise l'ordre de G. On a alors

$$|\mathcal{O}_x| = \frac{|G|}{|\mathrm{Stab}_G(x)|}.$$

Démonstration. L'application

$$\begin{aligned} f : G &\longrightarrow \mathcal{O}_x \\ g &\longmapsto g{\cdot}x \end{aligned}$$

est constante sur les classes à gauche modulo $\mathrm{Stab}_G(x)$.

En effet, si $s \in \mathrm{Stab}_G(x)$ et $g \in G$, on a alors

$$f(gs) = gs{\cdot}x = g{\cdot}(s{\cdot}x) = g{\cdot}x = f(g).$$

Comme l'ensemble des classes à gauche $G/\mathrm{Stab}_G(x)$ est l'ensemble des classes d'équivalence pour la relation associée au sous-groupe $\mathrm{Stab}_G(x)$, la proposition I-2.11 entraîne que f induit une application

$$\begin{aligned} \overline{f} : G/\mathrm{Stab}_G(x) &\longrightarrow \mathcal{O}_x \\ \overline{g} &\longmapsto g{\cdot}x. \end{aligned}$$

Par définition de l'orbite de x, l'application $\overline{f}$ est surjective. Montrons qu'elle est aussi injective. Soient $\overline{g}, \overline{g'} \in G/\mathrm{Stab}_G(x)$ tels que $g{\cdot}x = g'{\cdot}x$. On a alors

$$x = g^{-1}{\cdot}(g{\cdot}x) = g^{-1}{\cdot}(g'{\cdot}x) = g^{-1}g'{\cdot}x.$$

Ainsi, $g^{-1}g' \in \mathrm{Stab}_G(x)$ et donc $\overline{g} = \overline{g'}$. L'application $\overline{f}$ est donc bijective. Le reste de la proposition provient du théorème de Lagrange. □

On continue ce paragraphe en étudiant brièvement les orbites de E sous l'action d'un groupe monogène.

1.13. Définition. Soit G un groupe agissant sur un ensemble E.

Pour tout $g \in G$, et pour tout $x \in E$, la *g-orbite* de x est l'orbite de x sous l'action du sous-groupe $\langle g \rangle$.

1.14. Lemme. *Soit G un groupe agissant sur un ensemble E. Soit $g \in G$, et soit ω la g-orbite d'un élément $x \in E$. Alors, ω est finie si, et seulement si, il existe $k \in \mathbb{Z}$ non nul tel que $g^k \cdot x = x$. Dans ce cas, on a*

$$\omega = \{x,\ g \cdot x,\ \ldots,\ g^{p-1} \cdot x\},$$

où p est le plus petit entier strictement positif tel que $g^p \cdot x = x$. De plus, on a $|\omega| = p$.

Démonstration. Remarquons que par définition, on a

$$\omega = \{g^m \cdot x \mid m \in \mathbb{Z}\}.$$

Ce lemme peut se démontrer de la même façon que la première partie du théorème VI-4.13. Néanmoins, nous allons proposer ici une démonstration utilisant les actions de groupes.

Remarquons que l'application

$$\begin{aligned} \mathbb{Z} \times E &\longrightarrow E \\ (m, x') &\longmapsto g^m \cdot x' \end{aligned}$$

définit une action de $\mathbb{Z}$ sur E, et que l'orbite de x pour cette action est exactement ω. Par conséquent, d'après la proposition 1.12, ω est un ensemble fini si, et seulement si, $\mathrm{Stab}_{\mathbb{Z}}(x)$ est d'indice fini dans $\mathbb{Z}$. Or, $\mathrm{Stab}_{\mathbb{Z}}(x)$ est un sous-groupe de $\mathbb{Z}$. Par la proposition VI-2.4, il est donc de la forme $p\mathbb{Z}$, avec $p \geqslant 0$, et si $p > 0$, alors p est le plus petit élément strictement positif de $\mathrm{Stab}_{\mathbb{Z}}(x)$. D'après l'exemple VI-4.6, $p\mathbb{Z}$ est d'indice fini si, et seulement si, $p > 0$, i.e. si, et seulement si, $\mathrm{Stab}_{\mathbb{Z}}(x)$ est non trivial. Cela revient à dire qu'il existe $k \in \mathbb{Z}$ non nul tel que $g^k \cdot x = x$. Dans ce cas, on a

$$|\omega| = [\mathbb{Z} : \mathrm{Stab}_{\mathbb{Z}}(x)] = [\mathbb{Z} : p\mathbb{Z}] = p.$$

La bijection entre $\mathbb{Z}/p\mathbb{Z}$ et ω est alors donnée par

$$\begin{aligned} \mathbb{Z}/p\mathbb{Z} &\longrightarrow \omega \\ \overline{m} &\longmapsto g^m \cdot x, \end{aligned}$$

et comme les éléments distincts de $\mathbb{Z}/p\mathbb{Z}$ sont les classes $\overline{0}, \ldots, \overline{p-1}$, on a la conclusion. □

On finit en donnant une formule permettant parfois de faciliter le calcul du nombre d'orbites sous l'action d'un groupe fini G.

Notation. Soit G un groupe sur E. Pour tout $g \in G$, on pose

$$\mathrm{Fix}(g) = \{x \in E \mid g{\cdot}x = x\}.$$

Autrement dit, $\mathrm{Fix}(g)$ est l'ensemble des points fixes de la permutation

$$\begin{aligned} E &\longrightarrow E \\ x &\longmapsto g{\cdot}x. \end{aligned}$$

On a alors le résultat suivant.

1.15. Proposition (Formule de Burnside). *Soit G un groupe fini agissant sur un ensemble fini E. On note Ω l'ensemble des orbites de E sous l'action de G.*
Alors, on a l'égalité

$$|\Omega| = \frac{1}{|G|} \sum_{g \in G} |\mathrm{Fix}(g)|.$$

Démonstration. Soit

$$C = \{(g, x) \in G \times E \mid g{\cdot}x = x\}.$$

Calculons $|C|$ de deux manières différentes. Remarquons tout d'abord que l'on a

$$C = \bigcup_{x \in E} \{(g, x) \mid g \in G, g{\cdot}x = x\} = \bigcup_{x \in E} \mathrm{Stab}_G(x) \times \{x\}.$$

Cette réunion étant disjointe, on obtient

$$\begin{aligned} |C| &= \sum_{x \in E} |\mathrm{Stab}_G(x) \times \{x\}| \\ &= \sum_{x \in E} |\mathrm{Stab}_G(x)| \\ &= \sum_{x \in E} \frac{|G|}{|\mathcal{O}_x|} \\ &= |G| \sum_{x \in E} \frac{1}{|\mathcal{O}_x|}. \end{aligned}$$

Mais, E est la réunion disjointe de ses orbites sous l'action de G. Par ailleurs, si $\omega \in \Omega$, alors on a $\mathcal{O}_x = \omega$ pour tout $x \in \omega$. On a donc

$$\begin{aligned} |C| &= |G| \sum_{\omega \in \Omega} \sum_{x \in \omega} \frac{1}{|\mathcal{O}_x|} \\ &= |G| \sum_{\omega \in \Omega} \sum_{x \in \omega} \frac{1}{|\omega|} \\ &= |G| \sum_{\omega \in \Omega} |\omega| \frac{1}{|\omega|} \\ &= |G| \times |\Omega|. \end{aligned}$$

On peut également écrire C sous la forme de la réunion (disjointe) suivante :

$$C = \bigcup_{g\in G} \{(g,x) \mid x \in E,\ g\cdot x = x\} = \bigcup_{g\in G} \{g\} \times \mathrm{Fix}(g).$$

On a donc

$$|C| = \sum_{g\in G} |\mathrm{Fix}(g)|.$$

On en déduit la formule de Burnside

$$|\Omega| = \frac{1}{|G|} \sum_{g\in G} |\mathrm{Fix}(g)|.$$

Cela achève la démonstration. □

2. Premières applications

Nous allons maintenant donner quelques applications de la théorie des groupes opérants. D'autres seront proposées dans les exercices et dans les chapitres suivants.

On commence par un résultat dû à Cayley.

2.1. Théorème (Cayley). *Soit G un groupe fini d'ordre n. Alors, G est isomorphe à un sous-groupe de $\mathfrak{S}_n$.*

Démonstration. Pour le voir, faisons opérer G sur lui-même par translation à gauche. Cette action étant fidèle (cf. exemple 1.5 (1)), le morphisme de groupes correspondant

$$\varphi : G \longrightarrow \mathfrak{S}(G)$$

est injectif. Puisque G est d'ordre n, on a un isomorphisme [1]

$$\mathfrak{S}(G) \simeq \mathfrak{S}_n$$

et en composant par φ, on obtient un morphisme injectif $G \hookrightarrow \mathfrak{S}_n$, qui identifie donc G à son image dans $\mathfrak{S}_n$. □

Avant de donner d'autres applications, nous allons détailler un exemple d'action de groupe.

Soit p un nombre premier fixé et soit $n \geqslant 1$ un entier. Considérons G un groupe d'ordre n (il en existe au moins un, par exemple $G = \mathbb{Z}/n\mathbb{Z}$), et soit

$$E = \{x = (g_1, \ldots, g_p) \in G^p \mid g_1 \cdots g_p = 1_G\}.$$

1. Le lecteur pointilleux pourra, soit établir cet isomorphisme de lui-même, soit se référer à l'exemple VIII-1.3 du chapitre suivant.

C'est un ensemble fini, de cardinal n^{p-1}. En effet, l'ensemble E se récrit

$$E = \{x = (g_1, \dots, g_{p-1}, (g_1 \cdots g_{p-1})^{-1}) \mid g_1, \dots, g_{p-1} \in G\}.$$

Rappelons que $\mathfrak{S}_p$ agit sur G^p par

$$\begin{array}{ccc} \mathfrak{S}_p \times G^p & \longrightarrow & G^p \\ (\tau, (g_1, \dots, g_p)) & \longmapsto & (g_{\tau^{-1}(1)}, \dots, g_{\tau^{-1}(p)}). \end{array}$$

Soit $\sigma = (1\ p\ p-1 \cdots 2) \in \mathfrak{S}_p$, et soit $H = \langle \sigma \rangle$. Rappelons que l'on a

$$H = \{\mathrm{Id}, \sigma, \dots, \sigma^{p-1}\},$$

puisque σ est d'ordre p. Nous allons vérifier que l'action de $\mathfrak{S}_p$ sur G^p se restreint en une action de H sur E.

Notons tout d'abord que si $(g_1, g_2 \dots, g_p) \in E$, alors

$$(g_{\sigma^{-1}(1)}, \dots, g_{\sigma^{-1}(p)}) = (g_2, \dots, g_p, g_1) \in E.$$

En effet, si $g_1 g_2 \cdots g_p = 1_G$, alors

$$g_1^{-1}(g_1 g_2 \cdots g_p) g_1 = g_2 \cdots g_p g_1 = g_1^{-1} g_1 = 1_G.$$

Par récurrence, on obtient que si $(g_1, g_2 \dots, g_p) \in E$, alors

$$(g_{\sigma^{-m}(1)}, \dots, g_{\sigma^{-m}(p)}) \in E, \quad \text{pour tout } m \in [\![0, p-1]\!],$$

ce qu'il fallait vérifier.

Étudions les orbites de E sous l'action de H d'un peu plus près. D'après la proposition 1.12, le cardinal d'une orbite divise l'ordre de H, qui est p. Comme p est premier, on en déduit que le cardinal d'une orbite est 1 ou p. Décrivons maintenant les points fixes. Clairement, $x \in E^H$ si, et seulement si, toutes ses coordonnées sont égales.

Pour tout $x \in E$, on a donc $|\mathcal{O}_x| = 1$ si, et seulement si, on a $x = (g, \dots, g)$, où $g \in G$ vérifie $g^p = 1_G$. Remarquons que

$$g^p = 1_G \iff o(g) \mid p \iff o(g) = 1 \text{ ou } p.$$

Ainsi, $|\mathcal{O}_x| = 1$ si, et seulement si, $x = (g, \dots, g)$, avec $g = 1_G$ ou $o(g) = p$. Notons qu'il y a au moins une orbite réduite à un élément, à savoir l'orbite de $(1_G, \dots, 1_G)$.

Nous allons maintenant montrer l'équivalence des propriétés suivantes :

(1) on a $p \nmid n$;

(2) le groupe G ne possède pas d'élément d'ordre p ;

(3) il n'y a qu'une seule orbite réduite à un élément. Dans ce cas, on a la congruence $n^{p-1} \equiv 1 \ [p]$.

(1) $\Longrightarrow$ (2). Supposons que $p \nmid n$. Alors, G ne possède pas d'éléments d'ordre p. En effet, si un tel élément existait, son ordre diviserait l'ordre de G, qui est n, ce qui contredirait le choix de p.

(2) $\Longrightarrow$ (3). Si G ne possède pas d'éléments d'ordre p, il n'y a donc qu'une seule orbite réduite à un élément d'après l'analyse précédente. Mais alors, toutes les autres orbites ont p éléments. Comme les orbites distinctes forment une partition de E, le cardinal de E est la somme du cardinal des orbites[2]. En réduisant modulo p, on obtient $n^{p-1} \equiv 1 \ [p]$.

(3) $\Longrightarrow$ (1). Supposons qu'il y ait une unique orbite réduite à un élément. On a alors $n^{p-1} \equiv 1 \ [p]$. En particulier, on a $p \nmid n$.

Les trois propriétés sont donc bien équivalentes.

En appliquant (1) $\Longrightarrow$ (3), on retrouve le petit théorème de Fermat.

2.2. Théorème (Fermat). *Soit $n \geqslant 1$ un entier et soit p un nombre premier. Si $p \nmid n$, alors $n^{p-1} \equiv 1 \ [p]$.*

La contraposée de l'implication (2) $\Longrightarrow$ (1) nous donne aussi un premier résultat non trivial sur la structure des groupes finis.

2.3. Théorème (Cauchy). *Soit G un groupe fini et soit p un nombre premier. Si $p \mid |G|$, alors G possède au moins un élément d'ordre p.*

Continuons avec une autre application des actions de groupe. Le résultat suivant généralise le fait que tout sous-groupe d'indice 2 dans un groupe G est distingué dans G.

2.4. Théorème (Frobenius). *Soit G un groupe fini, et soit H un sous-groupe de G. Si tout diviseur premier de $|H|$ est $\geqslant [G : H]$, alors H est distingué dans G. En particulier, si p est le plus petit diviseur premier de $|G|$, alors tout sous-groupe d'indice p est distingué dans G.*

Démonstration. Pour appliquer la théorie des groupes opérants, nous allons retraduire la conclusion voulue en termes d'actions de groupe. On peut supposer que $H \neq G$, le cas $H = G$ étant trivial.

Soit $E = G/H$ l'ensemble des classes à gauche de G modulo H. Soit $h \in H$. Alors, on a les équivalences suivantes :

$$\begin{aligned} \text{pour tout } g \in G,\ ghg^{-1} \in H \quad &\Longleftrightarrow \quad \text{pour tout } g \in G,\ g^{-1}hg \in H \\ &\Longleftrightarrow \quad \text{pour tout } g \in G,\ hg \in gH \\ &\Longleftrightarrow \quad \text{pour tout } g \in G,\ \overline{hg} = \overline{g} \text{ dans } E. \end{aligned}$$

2. Plus rapidement, c'est simplement l'équation aux classes.

L'action de G sur E de l'exemple 1.5 (5) induit une action de H sur E par restriction, donnée par

$$H \times E \longrightarrow E$$
$$(h, \overline{g}) \longmapsto \overline{hg}.$$

Au vu des équivalences précédentes, H est distingué dans G si, et seulement si, tout élément de E est un point fixe sous l'action de H. Supposons donc qu'il existe $\overline{g} \in E$ qui ne soit pas un point fixe, et soit $\mathcal{O}_{\overline{g}}$ l'orbite correspondante. Alors, $\mathcal{O}_{\overline{g}}$ a au moins deux éléments.

Puisque E est réunion disjointe des orbites distinctes, on a en particulier

$$|\mathcal{O}_{\overline{g}}| \leqslant |E| = [G : H].$$

Soit ℓ un diviseur premier de $|\mathcal{O}_{\overline{g}}|$ (qui existe, car $|\mathcal{O}_{\overline{g}}| \geqslant 2$). Alors, d'après la proposition 1.12, on a

$$\ell \mid |\mathcal{O}_{\overline{g}}| \mid |H|.$$

Par hypothèse, on a alors

$$|\mathcal{O}_{\overline{g}}| \geqslant \ell \geqslant [G : H].$$

On a donc

$$|\mathcal{O}_{\overline{g}}| = [G : H] = |E|,$$

et en particulier $\mathcal{O}_{\overline{g}}$ est l'unique orbite de E sous l'action de G. On a donc

$$E = G/H = \mathcal{O}_{\overline{g}} = \mathcal{O}_{\overline{1}_G} = \{\overline{1}_G\},$$

d'où la contradiction $G = H$. Ainsi, tous les points de E sont fixes sous l'action de H, et H est distingué dans G. Le dernier point est une application directe du point précédent. □

Nous finissons ce paragraphe en donnant un résultat général sur l'existence de points fixes sous une action donnée. L'existence de tels points fixes n'est en général pas assurée. Par exemple, il suffit de considérer l'action d'un groupe G sur lui-même par translation à gauche.

En revanche, nous allons voir que la situation devient nettement plus sympathique pour une certaine classe de groupes finis.

2.5. Définition. Soit p un nombre premier. Un *p-groupe* est un groupe fini dont l'ordre est une puissance de p.

On a alors le résultat suivant.

2.6. Théorème. *Soit G un p-groupe agissant sur un ensemble fini E. Alors, on a*

$$|E| \equiv |E^G| \quad [p].$$

Si de plus $p \nmid |E|$, alors $E^G \neq \emptyset$.

Démonstration. On peut toujours supposer G non réduit à 1_G, sinon le résultat est clair (tout élément de E est fixé par 1_G). Posons $E' = E \setminus E^G$. Nous allons vérifier que E' est stable sous l'action de G. Cela revient à montrer que, pour tout $g \in G$ et tout $x \in E$, si $g{\cdot}x \in E^G$, alors $x \in E^G$.

Soient $g \in G$ et $x \in E$. Si $g{\cdot}x \in E^G$, on a en particulier

$$(gg'g^{-1}){\cdot}(g{\cdot}x) = g{\cdot}x, \quad \text{pour tout } g' \in G,$$

c'est-à-dire

$$gg'{\cdot}x = g{\cdot}x, \quad \text{pour tout } g' \in G.$$

En faisant agir g^{-1} des deux côtés de cette égalité, on en déduit

$$g'{\cdot}x = x, \quad \text{pour tout } g' \in G.$$

Ainsi, $x \in E^G$. L'ensemble E' est alors stable sous l'action de G. En particulier, cette action se restreint en une action sur E'.

Pour tout $x' \in E'$, on a $|\mathcal{O}_{x'}| \mid |G|$ par la proposition 1.12. De plus, l'orbite $\mathcal{O}_{x'}$ n'est pas réduite à un élément, car x' n'est pas un point fixe. Par conséquent, $|\mathcal{O}_{x'}|$ est une puissance non triviale de p (puisque $|G|$ est une puissance de p). Comme l'ensemble des orbites distinctes forment une partition de E', l'entier $|E'|$ est la somme des cardinaux de ces orbites, et l'on obtient donc

$$|E'| = |E| - |E^G| \equiv 0 \ [p],$$

soit encore

$$|E| \equiv |E^G| \ [p].$$

Supposons maintenant que $p \nmid |E|$. Si E^G était vide, on aurait $|E| \equiv 0 \ [p]$, d'après la congruence précédente, d'où une contradiction. □

On en déduit alors le résultat fort utile suivant.

2.7. Proposition. *Soit G un p-groupe non trivial. Alors, son centre $Z(G)$ est non trivial.*

Démonstration. On fait agir G sur lui-même par conjugaison. Alors, $Z(G)$ est exactement l'ensemble des points fixes pour cette action.

D'après le théorème 2.6, on a

$$|G| \equiv |Z(G)| \ [p].$$

Si l'on avait $Z(G) = \{1_G\}$, on obtiendrait

$$|G| \equiv 1 \ [p].$$

Mais, par hypothèse $p \mid |G|$, et l'on obtient une contradiction. □

3. Coloriages

Nous allons maintenant appliquer la théorie des actions de groupe au problème de coloriage. Imaginons un objet 2D ou 3D constitué d'un nombre fini de « parcelles » identiques indistinguables, que l'on peut colorier en utilisant un certain ensemble de couleurs (par exemple, une roulette, ou un collier de perles, ou un polyèdre régulier).

On voudrait compter le nombre de manières dont il est possible de colorier l'objet en utilisant un nombre donné de couleurs, ou encore le nombre de façons de colorier l'objet en utilisant chaque couleur à disposition un nombre de fois fixé à l'avance. Si l'on colorie l'objet en question, et qu'on le manipule sans le casser, on obtient toujours le même coloriage. Par exemple, si l'on considère une roulette à n secteurs, et que l'on lui imprime une rotation, le coloriage obtenu sera toujours le même. Plus généralement, on obtient toujours le même coloriage après avoir fait agir une isométrie de l'objet sur lui-même.

Les définitions suivantes formalisent cette situation.

3.1. Définition. Soit E un ensemble à n éléments. Soit $q \geqslant 1$ un entier. Un *coloriage* de E en au plus q couleurs est une application

$$f : E \longrightarrow [\![1, q]\!] .$$

On notera $\mathbf{C}(E, q)$ l'ensemble de ces coloriages. Si $f \in \mathbf{C}(E, q)$, son *type*, noté $\xi(f)$, est le q-uplet

$$\xi(f) = (n_1, \ldots, n_q) \in [\![0, n]\!]^q ,$$

où n_i est le nombre d'éléments $x \in E$ tels que $f(x) = i$ (c'est-à-dire le nombre de fois que l'on a utilisé la couleur i pour colorier notre objet).

Autrement dit, on a

$$n_i = |f^{-1}(\{i\})|.$$

Si $\xi \in [\![0, n]\!]^q$, on dira que f est *de type* ξ si $\xi(f) = \xi$.

Par exemple, si $E = \{x_1, x_2, x_3\}$ et $q = 4$, alors le coloriage $f : E \longrightarrow [\![1, q]\!]$ tel que

$$f(x_1) = f(x_2) = 1, \quad f(x_3) = 4$$

est de type $(2, 0, 0, 1)$.

On notera $\mathbf{C}(E, q, \xi)$ l'ensemble des coloriages de type ξ.

L'ensemble E représente en fait l'ensemble des « parcelles » de l'objet à colorier. Nous voulons maintenant traduire le fait que l'on veut identifier deux coloriages après avoir fait agir une isométrie de notre objet sur lui-même. En fait, on va se placer dans un cadre un tout petit peu plus général.

Soit G un groupe (fini ou non) agissant sur E. On note $\varphi : G \longrightarrow \mathfrak{S}(E)$ le morphisme de groupes correspondant. Pour tout $g \in G$, et tout $f \in \mathbf{C}(E,q)$, on note $g{\cdot}f$ l'élément de $\mathbf{C}(E,q)$ défini par

$$(g{\cdot}f)(x) = f(g^{-1}{\cdot}x) = f\big(\varphi(g)^{-1}(x)\big), \quad \text{pour tout } x \in E.$$

3.2. Lemme. *L'application*

$$\begin{array}{ccc} G \times \mathbf{C}(E,q) & \longrightarrow & \mathbf{C}(E,q) \\ (g,f) & \longmapsto & g{\cdot}f \end{array}$$

définit une action de G sur $\mathbf{C}(E,q)$. De plus, elle induit par restriction une action de G sur $\mathbf{C}(E,q,\xi)$.

Démonstration. Pour tout $x \in E$, on a

$$(1_G{\cdot}f)(x) = f(1_G^{-1}{\cdot}x) = f(x),$$

donc $1_G{\cdot}f = f$.

Soient $g, g' \in G$, et soit $f \in \mathbf{C}(E,q)$. Pour tout $x \in E$, on a

$$\begin{array}{rcl} \big(g{\cdot}(g'{\cdot}f)\big)(x) & = & (g'{\cdot}f)\big(g^{-1}(x)\big) \\ & = & f\big(g'^{-1}\big(g^{-1}(x)\big)\big) \\ & = & f\big((gg')^{-1}(x)\big) \\ & = & (gg'{\cdot}f)(x), \end{array}$$

donc $gg'{\cdot}f = g{\cdot}(g'{\cdot}f)$, et l'on a bien une action de G sur $\mathbf{C}(E,q)$.

Remarquons de plus que pour tout $g \in G$, et tout $f \in \mathbf{C}(E,q)$, on a

$$\xi(f) = \xi(g{\cdot}f).$$

En effet, pour tout $x \in E$ et tout $i \in [\![1,q]\!]$, on a

$$(g{\cdot}f)(x) = i \iff f(g^{-1}{\cdot}x) = i.$$

On a donc

$$x \in (g{\cdot}f)^{-1}(\{i\}) \iff g^{-1}{\cdot}x \in f^{-1}(\{i\}).$$

L'application

$$\begin{array}{rccc} \sigma_{g^{-1}} : & E & \longrightarrow & E \\ & x & \longmapsto & g^{-1}{\cdot}x \end{array}$$

étant bijective, l'équivalence précédente montre que

$$|(g{\cdot}f)^{-1}(\{i\})| = |f^{-1}(\{i\}|, \quad \text{pour tout } i \in [\![1,q]\!],$$

ce qu'il fallait vérifier. On peut donc restreindre cette action en une action de G sur $\mathbf{C}(E,q,\xi)$. □

3.3. Définition. On dit que deux coloriages $f, f' \in \mathbf{C}(E,q)$ (resp. deux coloriages $f, f' \in \mathbf{C}(E,q,\xi)$) sont *$G$-équivalents* s'ils sont dans la même orbite sous l'action de G, c'est-à-dire s'il existe $g \in G$ tel que $f' = g{\cdot}f$. On note $\mathbf{C}_G(E,q)$ (resp. $\mathbf{C}_G(E,q,\xi)$) l'ensemble des orbites des coloriages (resp. l'ensemble des orbites des coloriages de type ξ) sous l'action de G.

Cela correspond bien à l'idée intuitive que l'on se fait de l'identification de deux coloriages. Si G est le groupe des isométries de l'objet à colorier, cette définition dit simplement que deux coloriages sont équivalents si l'on peut passer de l'un à l'autre en manipulant l'objet en question.

Le but de ce qui suit est de calculer le nombre de coloriages de type donné deux à deux non équivalents, ainsi que le nombre total de coloriages deux à deux non équivalents sous l'action de G, c'est-à-dire le nombre d'éléments de $\mathbf{C}_G(E,q,\xi)$ et $\mathbf{C}_G(E,q)$.

Rappelons que l'action de G sur $\mathbf{C}(E,q,\xi)$ est définie de la manière suivante : pour tout $g \in G$, tout $f \in \mathbf{C}(E,q,\xi)$ et tout $x \in E$, on a

$$(g{\cdot}f)(x) = f(g^{-1}{\cdot}x) = f\big(\varphi(g)^{-1}(x)\big),$$

où $\varphi : G \longrightarrow \mathfrak{S}(E)$ est le morphisme de groupes induit par l'action de G sur E.

Remarquons qu'un sous-groupe S de $\mathfrak{S}(E)$ agit de façon naturelle sur E par

$$\begin{aligned} S \times E &\longrightarrow E \\ (\sigma, x) &\longmapsto \sigma(x), \end{aligned}$$

et induit donc une action $*$ sur $\mathbf{C}(E,q,\xi)$, qui est simplement définie par

$$(\sigma * f)(x) = f\big(\sigma^{-1}(x)\big), \quad \text{pour tout } f \in \mathbf{C}(E,q,\xi), \text{ et tout } x \in E.$$

En particulier, c'est le cas pour $S = \varphi(G)$. Mais alors, pour tout $g \in G$, et tout $f \in \mathbf{C}(E,q,\xi)$, on a $g{\cdot}f = \varphi(g) * f$. Ainsi, le nombre d'orbites de $\mathbf{C}(E,q,\xi)$ sous l'action de G et le nombre d'orbites de $\mathbf{C}(E,q,\xi)$ sous l'action de $\varphi(G)$ sont égaux.

Quitte à remplacer G par $\varphi(G)$, on peut donc supposer que G est fini. Bien que nous ne le ferons pas dans la suite, on peut même supposer si nécessaire que G est un sous-groupe de $\mathfrak{S}(E)$, et que G agit naturellement sur E.

On suppose dorénavant que G est un groupe fini. Nous allons maintenant définir un polynôme à plusieurs variables qui va contenir toutes les informations sur les coloriages.

Notation. Soient $X_1, \ldots, X_q$ q indéterminées. Si $\xi = (n_1, \ldots, n_q) \in \mathbb{N}^q$, on pose $X^\xi = X_1^{n_1} \cdots X_q^{n_q} \in \mathbb{Q}[X_1, \ldots, X_q]$.

3.4. Définition. Soit E un ensemble fini, et soit G un groupe fini agissant sur E.

On appelle *inventaire* de E sous l'action de G le polynôme

$$W_{E,G} = \sum_{\xi \in [\![0,n]\!]^q} |\mathbf{C}_G(E,q,\xi)| X^\xi \in \mathbb{Q}[X_1,\ldots,X_q].$$

3.5. Lemme. *On a*

$$W_{E,G} = \frac{1}{|G|} \sum_{g\in G} \sum_{f \in \mathbf{C}(E,q)(g)} X^{\xi(f)},$$

où $\mathbf{C}(E,q)(g) = \{f \in \mathbf{C}(E,q) \mid g{\cdot}f = f\}$.

Démonstration. Une application directe de la formule de Burnside montre que l'on a

$$|\mathbf{C}_G(E,q,\xi)| = \frac{1}{|G|} \sum_{g\in G} |\mathbf{C}(E,q,\xi)(g)|,$$

où $\mathbf{C}(E,q,\xi)(g) = \{f \in \mathbf{C}(E,q,\xi) \mid g{\cdot}f = f\}$. Sommons cette égalité sur l'ensemble des types possibles.

On obtient alors

$$\sum_{\xi \in [\![0,n]\!]^q} |\mathbf{C}_G(E,q,\xi)| X^\xi = W_{E,G} = \frac{1}{|G|} \sum_{\xi \in [\![0,n]\!]^q} \sum_{g\in G} |\mathbf{C}(E,q,\xi)(g)| X^\xi.$$

Comme ξ et g sont indépendants, on a

$$W_{E,G} = \frac{1}{|G|} \sum_{g\in G} \sum_{\xi \in [\![0,n]\!]^q} |\mathbf{C}(E,q,\xi)(g)| X^\xi.$$

Si $g \in G$, alors $\mathbf{C}(E,q)(g)$ est la réunion disjointe des $\mathbf{C}(E,q,\xi)(g)$, lorsque ξ parcourt $[\![0,n]\!]^q$.

Comme tout élément $f \in \mathbf{C}(E,q,\xi)(g)$ est de type ξ par définition, on obtient donc

$$\begin{aligned} \sum_{f\in \mathbf{C}(E,q)(g)} X^{\xi(f)} &= \sum_{\xi \in [\![0,n]\!]^q} \sum_{f \in \mathbf{C}(E,q,\xi)(g)} X^{\xi(f)} \\ &= \sum_{\xi \in [\![0,n]\!]^q} \sum_{f \in \mathbf{C}(E,q,\xi)(g)} X^{\xi} \\ &= \sum_{\xi \in [\![0,n]\!]^q} |\mathbf{C}(E,q,\xi)(g)| X^\xi. \end{aligned}$$

On en déduit ainsi l'égalité

$$W_{E,G} = \frac{1}{|G|} \sum_{g\in G} \sum_{f \in \mathbf{C}(E,q)(g)} X^{\xi(f)}.$$

□

Le but est maintenant de trouver une expression manipulable de $W_{E,G}$. On commence par étudier les éléments de $\mathbf{C}(E,q)(g)$.

3.6. Lemme. *Soit $g \in G$. Alors, un coloriage f est fixé par g si, et seulement si, f est constant sur chaque g-orbite.*

Démonstration. Supposons que le coloriage f soit fixé par g, c'est-à-dire

$$f(g^{-1}\cdot x) = f(x), \qquad \text{pour tout } x \in E.$$

En remplaçant x par $g\cdot x$, cela se récrit

$$f(g\cdot x) = f(x), \qquad \text{pour tout } x \in E.$$

Si ω est la g-orbite de x, alors on a $\omega = \{g^m\cdot x \mid m \in \mathbb{Z}\}$. Les relations précédentes et une récurrence facile montrent alors que la fonction f est constante sur ω.

Inversement, supposons que f soit constante sur les g-orbites et soit $x \in E$. Puisque $g^{-1}\cdot x$ et x sont dans la même g-orbite, on a

$$f(g^{-1}\cdot x) = f(x), \qquad \text{pour tout } x \in E,$$

et donc $g\cdot f = f$. □

Soit $g \in G$, et soient $\omega_1, \ldots, \omega_r$ les diverses g-orbites de E. Soit $Y = [\![1, r]\!]$. On a la bijection suivante entre $\mathbf{C}(Y,q)$ et $\mathbf{C}(E,q)(g)$:

si $h \in \mathbf{C}(Y,q)$, on associe le coloriage $\theta(h) \in \mathbf{C}(E,q)(g)$ défini par

$$\theta(h)(x) = h(i),$$

où ω_i est la g-orbite de x. Par construction, le coloriage $\theta(h)$ est constant sur les g-orbites, et donc fixé par g par le lemme précédent.

Réciproquement, si $f \in \mathbf{C}(E,q)(g)$, on définit un coloriage $\rho(f) \in \mathbf{C}(Y,q)$ par la formule

$$\rho(f)(i) = f(x),$$

où x est un élément arbitraire de ω_i. Ce coloriage est bien défini (i.e. ne dépend pas du choix de x), puisque f est constant sur les g-orbites.

On vérifie aisément que θ et ρ sont inverses l'une de l'autre.

3.7. Lemme. *Soit $g \in G$, et soient $p_1, \ldots, p_r$ les cardinaux respectifs des diverses g-orbites $\omega_1, \ldots, \omega_r$ de E. Alors, on a*

$$\sum_{f \in \mathbf{C}(E,q)(g)} X^{\xi(f)} = \prod_{i=1}^{r} (X_1^{p_i} + \cdots + X_q^{p_i}).$$

Démonstration. On remarque d'abord que si $f : E \longrightarrow [\![1, q]\!]$ est un colo-

riage arbitraire, alors on a $X^{\xi(f)} = \prod_{x\in E} X_{f(x)}$. Maintenant, puisque l'application $\theta : \mathbf{C}(Y,q) \longrightarrow \mathbf{C}(E,q)(g)$ est une bijection, on a

$$\sum_{f\in\mathbf{C}(E,q)(g)} X^{\xi(f)} = \sum_{h\in\mathbf{C}(Y,q)} X^{\xi(\theta(h))} = \sum_{h\in\mathbf{C}(Y,q)} \prod_{x\in E} X_{\theta(h)(x)}.$$

Comme E est la réunion disjointe de $\omega_1,\dots,\omega_r$, et puisque $\theta(h)$ vaut $h(i)$ sur chaque ω_i, on a

$$\sum_{f\in\mathbf{C}(E,q)(g)} X^{\xi(f)} = \sum_{h\in\mathbf{C}(Y,q)} \prod_{i=1}^{r} X_{h(i)}^{p_i}.$$

En remarquant que $h(i)$ peut prendre toute les valeurs $m_i \in [\![1,q]\!]$, on en déduit l'égalité

$$\sum_{f\in\mathbf{C}(E,q)(g)} X^{\xi(f)} = \sum_{m_1,\dots,m_r\in[\![1,q]\!]} \prod_{i=1}^{r} X_{m_i}^{p_i} = \sum_{m_1,\dots,m_r\in[\![1,q]\!]} X_{m_1}^{p_1}\cdots X_{m_r}^{p_r}.$$

On reconnaît alors le terme général du produit $\prod_{i=1}^{r}(X_1^{p_i}+\cdots+X_q^{p_i})$. En effet, pour développer ce produit, on choisit un terme de la somme dans chaque facteur et l'élément correspondant est alors de la forme

$$X_{m_1}^{p_1}\cdots X_{m_r}^{p_r}.$$

Ainsi, on a bien

$$\sum_{f\in\mathbf{C}(E,q)(g)} X^{\xi(f)} = \prod_{i=1}^{r}(X_1^{p_i}+\cdots+X_q^{p_i}). \qquad \square$$

3.8. Théorème (Pólya). *Soit E un ensemble à n éléments, et soit G un groupe fini agissant sur E. On a*

$$W_{E,G} = \frac{1}{|G|}\sum_{g\in G}\prod_{i=1}^{n}(X_1^i+\cdots+X_q^i)^{e_i(g)},$$

où $e_i(g)$ est le nombre de g-orbites à i éléments.

Démonstration. D'après le lemme précédent, on a

$$\sum_{f\in\mathbf{C}(E,q)(g)} X^{\xi(f)} = \prod_{i=1}^{r}(X_1^{p_i}+\cdots+X_q^{p_i}),$$

où $p_1,\dots,p_r$ sont les cardinaux des diverses g-orbites. Puisque le cardinal d'une g-orbite est toujours compris entre 1 et n, cela peut se récrire en regroupant les facteurs identiques

$$\sum_{f\in\mathbf{C}(E,q)(g)} X^{\xi(f)} = \prod_{k=1}^{n}(X_1^k+\cdots+X_q^k)^{e_k(g)}.$$

On utilise alors le lemme 3.5 pour conclure. $\square$

3.9. Corollaire. *Le nombre $|\mathbf{C}_G(E,q)|$ de classes de G-équivalence de coloriages de E en au plus q couleurs est égal à*

$$|\mathbf{C}_G(E,q)| = \frac{1}{|G|}\sum_{g\in G} q^{r(g)},$$

où $r(g)$ est le nombre de g-orbites.

Démonstration. On a

$$|\mathbf{C}_G(E,q)| = \sum_{\xi\in[\![0,n]\!]^q} |\mathbf{C}_G(E,q,\xi)| = W_{E,G}(1,\ldots,1).$$

On utilise alors le théorème précédent pour conclure, en remarquant que l'on a $r(g) = \sum_{i=1}^{n} e_i(g)$. □

3.10. Remarque. Le lecteur peu à l'aise avec les polynômes en plusieurs indéterminées, et seulement intéressé par le problème de dénombrement de tous les coloriages, pourra remarquer que ce dernier corollaire s'obtient directement en appliquant la formule de Burnside et le lemme 3.6.

Application au problème de la roulette

On considère une roulette à n secteurs, et l'on dispose de q couleurs pour la colorier. On identifie deux coloriages si l'on peut passer de l'une à l'autre par une rotation.

Question 1. Combien de coloriages différents y a-t-il ?

On identifie chaque secteur de la roulette au sommet d'un n-gone régulier, que l'on numérote de 1 à n dans le sens trigonométrique, et l'on considère le groupe cyclique G d'ordre n engendré par la rotation r de centre O (le centre du n-gone) et d'angle de mesure $\frac{2\pi}{n}$. Alors, G agit naturellement sur l'ensemble E des sommets du n-gone.

Tourner la roulette revenant à faire agir G sur E, on doit donc calculer le cardinal de $\mathbf{C}_G(E,q)$, où $G = \langle r\rangle$ (deux coloriages sont identifiés si, et seulement si, ils sont G-équivalents au sens de la section précédente).

Il s'agit donc calculer le nombre de g-orbites de longueur fixée pour tout élément $g \in G$, d'après le corollaire 3.9.

Rappelons qu'une g-orbite n'est rien d'autre qu'une orbite de E sous l'action du groupe $\langle g\rangle$. Le cardinal de l'orbite d'un sommet S sous cette action est donc donné par

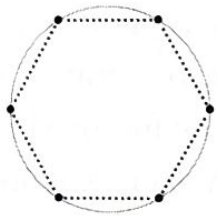
$$\frac{|\langle g\rangle|}{|\mathrm{Stab}_{\langle g\rangle}(S)|}.$$

Par définition de l'ordre d'un élément, on a $|\langle g\rangle| = o(g)$.

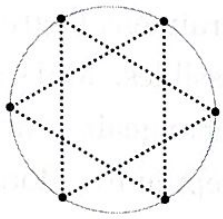
Mais, l'unique rotation du plan de centre O possédant un point fixe différent de O est l'identité. Par conséquent, $|\mathrm{Stab}_{\langle g\rangle}(S)| = 1$.

Ainsi, les g-orbites sont toutes de longueur $o(g)$ et il y en a donc $\dfrac{n}{o(g)}$. Cet entier ne dépend donc que de l'ordre de $g \in G$, qui est un diviseur de n.

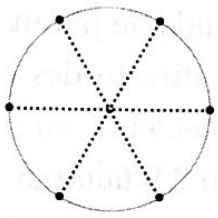
La figure suivante illustre le cas $n = 6$. On a donc $d \in \{1, 2, 3, 6\}$.



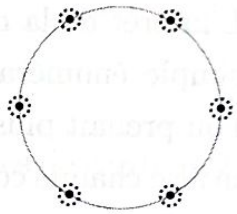
Rotation d'ordre $d = 6$
$\frac{n}{d} = 1$ orbite



Rotation d'ordre $d = 3$
$\frac{n}{d} = 2$ orbites



Rotation d'ordre $d = 2$
$\frac{n}{d} = 3$ orbites



Identité, d'ordre $d = 1$
$\frac{n}{d} = 6$ orbites

Le nombre de coloriages différents d'une roulette à n secteurs en au plus q couleurs est donc donné par la formule

$$\frac{1}{n}\sum_{d|n}\varphi(d)q^{\frac{n}{d}},$$

où l'entier $\varphi(d)$ est le nombre d'éléments d'ordre d dans un groupe cyclique d'ordre n (cf. corollaire VI-6.7 pour une formule explicite).

Remarquons que l'inventaire est dans ce cas

$$W_{E,G} = \frac{1}{n}\sum_{d|n}\varphi(d)(X_1^d + \cdots + X_q^d)^{\frac{n}{d}}.$$

Question 2. Prenons une roulette à six secteurs. Combien de fois peut-on colorier cette roulette en utilisant trois fois le rouge, deux fois le vert, et une fois le bleu ?

Disons que le rouge est la couleur 1, le vert la couleur 2 et le bleu la couleur 3. En reprenant les notations de la section précédente, on cherche donc $|\mathbf{C}_G(E, q, \xi)|$, où $\xi = (3, 2, 1)$. D'après la définition de l'inventaire, l'entier $|\mathbf{C}_G(E, q, \xi)|$ est le coefficient du monôme $X^\xi = X_1^3X_2^2X_3$. D'après le théorème de Polya, $|\mathbf{C}_G(E, q, \xi)|$ est donc le coefficient du monôme X^ξ dans le polynôme

$$\frac{1}{6}\sum_{g\in G}\prod_{i=1}^{6}(X_1^i + X_2^i + X_3^i)^{e_i(g)},$$

lequel est égal à

$$\frac{1}{6}\sum_{d|6}\varphi(d)(X_1^d + X_2^d + X_3^d)^{\frac{6}{d}},$$

ou encore

$$\frac{1}{6}\big((X_1+X_2+X_3)^6+(X_1^2+X_2^2+X_3^2)^3\big)+\frac{1}{3}\big((X_1^3+X_2^3+X_3^3)^2+(X_1^6+X_2^6+X_3^6)\big).$$

Par conséquent, $|\mathbf{C}_G(E,q,\xi)| = \frac{1}{6}\binom{6}{3}\binom{3}{1} = 10$. Il y a donc dix roulettes différentes à six secteurs dont trois rouges, deux verts et un bleu.

L'intérêt de la méthode ne paraît pas vraiment flagrant par rapport à une simple énumération directe des cas possibles. Mais que se passerait-il si l'on prenait plus de couleurs, ou si l'on changeait le nombre de fois que l'on utilise chaque couleur ? Il faudrait tout reprendre, alors qu'avec l'inventaire, il suffit de calculer le coefficient du monôme adéquat.

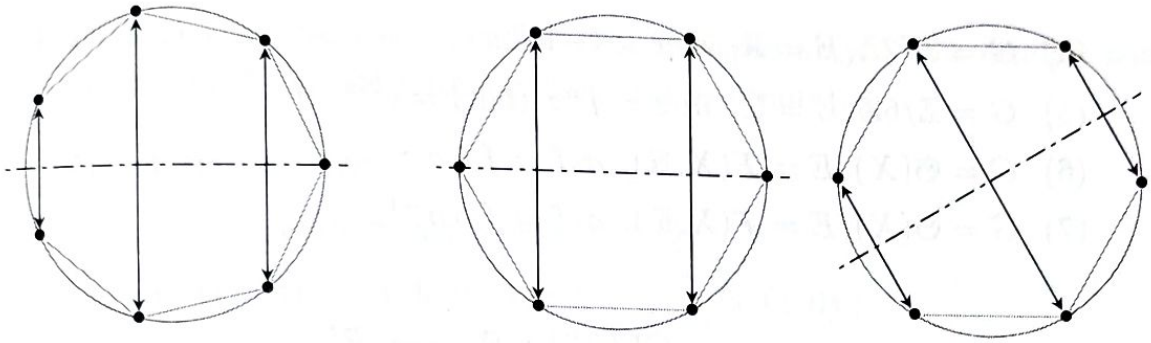
Application au problème du collier

On considère cette fois un collier de n perles de couleurs de même taille (numérotées de 1 à n). On dispose de q couleurs différentes. On identifie deux colliers si l'on peut les obtenir après rotation ou en les retournant. En identifiant les perles aux sommets d'un n-gone régulier, on constate que cette fois, c'est le groupe diédral D_n qui agit sur l'ensemble des sommets du polygone. Il est constitué des rotations de centre O (le centre du n-gone) et d'angle de mesure $\frac{2k\pi}{n}$, avec $k \in [\![0, n-1]\!]$, et de n symétries orthogonales, dont les axes passent par O.

Comme précédemment, le nombre d'orbites d'une rotation d'ordre d (où d divise n) est $\frac{n}{d}$. Pour les symétries orthogonales, il faut distinguer selon la parité de n.

(a) Supposons que $n = 2m-1, m \geqslant 1$, et soit s une symétrie orthogonale. Dans ce cas, l'axe de s passe par un sommet du n-gone et un seul. On voit facilement que s fixe le sommet appartenant à l'axe et échange les autres sommets par paire. Il y a alors 1 orbite de cardinal 1, et $m-1$ orbites de cardinal 2, soit m orbites pour chacune des n symétries.

(b) Si $n = 2m$, il y a m symétries orthogonales dont l'axe passe par deux sommets, et m symétries orthogonales dont l'axe ne passe par aucun sommet. Dans le premier cas, la symétrie échange les $n-2$ sommets n'appartenant pas à l'axe par paire, et a deux points fixes. Dans le second cas, la symétrie échange les n sommets par paire, et n'a aucun point fixe.



n impair n pair

Ainsi, on a

• m symétries, avec deux orbites de cardinal 1 et $m-1$ orbites de cardinal 2, soit $m+1$ orbites.

• m symétries, avec m orbites de cardinal 2.

Si n est impair, on a alors

$$|\mathbf{C}_G(E,q)| = \frac{1}{2n}\Big(\sum_{d|n}\varphi(d)q^{\frac{n}{d}} + nq^{\frac{n+1}{2}}\Big).$$

Si n est pair, on a alors

$$|\mathbf{C}_G(E,q)| = \frac{1}{2n}\Big(\sum_{d|n}\varphi(d)q^{\frac{n}{d}} + \frac{n}{2}q^{\frac{n+2}{2}} + \frac{n}{2}q^{\frac{n}{2}}\Big).$$

On laisse le soin au lecteur d'écrire l'inventaire dans chacun des cas.

Les considérations sur les coloriages et les arguments présentés ici sont ceux de G. Pólya. Néanmoins, nous nous sommes très largement inspiré de l'article de Thomas Chomette [16].

4. Exercices

E1. Soit G un groupe agissant sur un ensemble E. Démontrer en détail les résultats exposés dans la remarque 1.2.

E2. Détailler les exemples du point 1.5.

E3. Dans les cas suivants, l'application $\begin{array}{c} G\times E \longrightarrow E \\ (g,x)\longmapsto g\cdot x \end{array}$ est-elle une action de G sur E ?

(1) $G=\mathbb{Z}, E=\mathbb{R}$, $n\cdot x = x+n$;

(2) $G=\mathbb{Z}, E=\mathbb{R}$, $n\cdot x = nx$;

(3) $G=\mathbb{Z}, E=\mathbb{R}$, $n\cdot x = 2^n x$;

(4) $G = \mathbb{Z}/2\mathbb{Z}, E = \mathbb{R}$, $\overline{m}\cdot x = (-1)^m x$;

(5) $G = \mathbb{Z}/6\mathbb{Z}, E = \mathbb{C}$, $\overline{m}\cdot z = j^m z$ (où $j = e^{2i\pi/3}$) ;

(6) $G = \mathfrak{S}(X), E = \mathcal{F}(X, \mathbb{R})$, $\sigma\cdot f = f \circ \sigma$;

(7) $G = \mathfrak{S}(X), E = \mathcal{F}(X, \mathbb{R})$, $\sigma\cdot f = f \circ \sigma^{-1}$.

E4. On considère l'application
$$\begin{aligned} \mathrm{GL}(\mathbb{R}^n) \times \mathbb{R}^n &\longrightarrow \mathbb{R}^n \\ (f, x) &\longmapsto f(x). \end{aligned}$$

a. Montrer que l'on définit ainsi une action de $\mathrm{GL}(\mathbb{R}^n)$ sur $\mathbb{R}^n$.

b. Cette action est-elle transitive ? Décrire les orbites.

c. Pour $n = 2$, décrire le stabilisateur de $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$, puis de $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

E5. Soit $E = \mathbb{C} \setminus \{-1, 0, 1\}$, et soit
$$\begin{aligned} \sigma : E &\longrightarrow \mathbb{C} \\ z &\longmapsto \frac{z-1}{z+1}. \end{aligned}$$

a. Montrer que $\sigma \in \mathfrak{S}(E)$.

b. Déterminer le cardinal de $G = \langle\sigma\rangle$.

c. Justifier que l'application $\begin{aligned} G \times E &\longrightarrow E \\ (g, x) &\longmapsto g(x) \end{aligned}$ définit une action de G sur E. Déterminer l'orbite et le stabilisateur de chaque élément de E.

E6. Soit G un groupe agissant sur un ensemble E.
Si $X \subset E$, on note $g\cdot X = \{g\cdot x \mid x \in X\}$. Le stabilisateur de X sous l'action de G est l'ensemble
$$\mathrm{Stab}_G(X) = \{g \in G \mid g\cdot X = X\}.$$

a. Montrer que $\mathrm{Stab}_G(X)$ est un sous-groupe de G, et montrer que l'action de G sur E se restreint en une action de $\mathrm{Stab}_G(X)$ sur X.

b. Montrer que pour tout $g_0 \in G$, on a
$$\mathrm{Stab}_G(g_0\cdot X) = g_0 \mathrm{Stab}_G(X) g_0^{-1}.$$
En déduire que deux éléments de E dans la même G-orbite ont des stabilisateurs isomorphes.

On suppose jusqu'à la fin que $G = \mathrm{GL}(\mathbb{R}^2)$ et $E = \mathbb{R}^2$.

c. Si $X \subset E$ est fini, $\mathrm{Stab}_G(X)$ est-il fini ?

d. Dans chacun des cas suivants, déterminer $\mathrm{Stab}_G(X)$ et étudier son action sur X (transitivité, orbites, stabilisateurs).

(1) X est l'axe Ox ;

(2) X est la réunion des axes Ox et Oy ;

(3) X est formé des deux points $(-1,0)$ et $(1,0)$;

(4) X est l'ensemble des sommets du triangle équilatéral

$$A=(1,0), \quad B=\left(\cos\left(\frac{2\pi}{3}\right),\sin\left(\frac{2\pi}{3}\right)\right), \quad C=\left(\cos\left(\frac{4\pi}{3}\right),\sin\left(\frac{4\pi}{3}\right)\right);$$

(5) X est l'ensemble des sommets du rectangle

$$A=(2,1), \quad B=(-2,1), \quad C=(-2,-1), \quad D=(2,-1).$$

E7. Étudier l'action par conjugaison de $\mathrm{O}(\mathbb{R}^2)$ sur lui-même (orbites, points fixes, stabilisateurs).

E8. Soit G un groupe agissant sur un ensemble E.

a. On suppose que $|G| = 156$, et qu'il existe un $x \in E$ avec un stabilisateur d'ordre 12. Quel est le cardinal de l'orbite de x ?

b. On suppose que $\mathrm{card}(E) = 108$ et $|G| = 143$. Montrer que G fixe au moins un point de E.

c. On suppose que $\mathrm{card}(E) = 17$ et $|G| = 15$. On suppose que G agit sans point fixe. Déterminer le nombre d'orbites et leurs cardinaux.

d. On suppose que $\mathrm{card}(E) = 27$ et $|G| = 63$. L'action de G peut-elle être transitive ?

E9. Soit K un corps. Le groupe $\mathrm{GL}_2(K)$ agit naturellement sur K^2 par multiplication matricielle.

a. Combien y a-t-il d'orbites ?

b. On suppose que $K = \mathbb{F}_2$. Combien y a-t-il d'éléments non nuls dans K^2 ? Montrer que $\mathrm{GL}_2(K)$ agit naturellement sur un ensemble à trois éléments, de façon à ce que le morphisme induit $\mathrm{GL}_2(K) \longrightarrow \mathfrak{S}_3$ soit un isomorphisme de groupes.

E10. Soit p un nombre premier, et soit G un groupe fini d'ordre p^2.

a. Justifier que $Z(G)$ est d'ordre p ou p^2.

b. On suppose que $|Z(G)| = p$. On fait agir G sur lui-même par conjugaison. Soit $x \in G \setminus Z(G)$.

(i) Montrer que $x \in \mathrm{Stab}_G(x)$, puis que $Z(G) \subsetneq \mathrm{Stab}_G(x)$.

(ii) En déduire que $\mathrm{Stab}_G(x) = G$, puis une contradiction.

c. En déduire que G est commutatif.

d. Montrer que G est isomorphe à $\mathbb{Z}/p^2\mathbb{Z}$ ou $\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}$.

Indication. Utiliser l'exercice VI-43.

E11. Soit G un groupe, et soient X, Y deux ensembles non vides sur lesquels G agit. On dit qu'une application $f : X \longrightarrow Y$ est *G-équivariante* si l'on a

$$f(g{\cdot}x) = g{\cdot}f(x), \quad \text{pour tout } g \in G \text{ et tout } x \in X.$$

a. Soit $f : X \longrightarrow Y$ une application G-équivariante.

(i) Montrer que f induit une application $f_* : X^G \longrightarrow Y^G$.

(ii) On suppose que f est injective. Montrer que f_* est injective.

(iii) On suppose que f est bijective. Montrer que f_* est bijective.

(iv) On suppose que f est surjective. L'application f_* est-elle nécessairement surjective ?

On dit que les ensembles X et Y sont *G-isomorphes* s'il existe une bijection $f : X \xrightarrow{\sim} Y$ G-équivariante.

b. Montrer que tout ensemble non vide sur lequel G agit transitivement est G-isomorphe à un ensemble de la forme G/H, où H est un sous-groupe de G, et où G agit sur G/H par translation à gauche.

E12. Soit G un groupe fini, soit p le plus petit diviseur premier de $|G|$ et soit H un sous-groupe distingué de G d'ordre p.
Montrer que $H \subset Z(G)$.

Indication. Faire opérer G sur H par conjugaison.

E13. Soit G un groupe infini. On suppose que G admet un sous-groupe H, distinct de G et d'indice fini. Montrer que G n'est pas simple.

Indication. Faire opérer G sur G/H par translation.

Le résultat de l'exercice suivant est tiré de l'excellent article de Keith Conrad [21].

E14. Soit G un groupe fini. Le groupe $\mathrm{Aut}(G)$ agit naturellement sur G par

$$\begin{aligned} \mathrm{Aut}(G) \times G &\longrightarrow G \\ (\varphi, g) &\longmapsto \varphi(g). \end{aligned}$$

a. Justifier que l'action de $\mathrm{Aut}(G)$ sur G se restreint en une action sur l'ensemble $G \setminus \{1_G\}$.

b. On suppose que $G \simeq (\mathbb{Z}/p\mathbb{Z})^m$, où p est premier et $m \geqslant 1$. Montrer que $\mathrm{Aut}(G)$ agit transitivement sur $G \setminus \{1_G\}$.

On suppose maintenant que G est un groupe fini tel que $\mathrm{Aut}(G)$ agisse transitivement sur $G \setminus \{1_G\}$.

c. Montrer que tous les éléments de $G \setminus \{1_G\}$ ont même ordre.

d. En déduire qu'il existe un nombre premier p et un entier $m \geqslant 1$ tel que G soit d'ordre p^m, et que tous les éléments de $G \setminus \{1_G\}$ sont d'ordre p.

e. Justifier l'existence d'un élément de $z \in Z(G) \setminus \{1_G\}$. Montrer que G est abélien.

Indication. Soient $\varphi \in \mathrm{Aut}(G)$ et $z \in G$ tels que $\varphi(z) \in Z(G)$. Que dire de z ?

f. En déduire que $G \simeq (\mathbb{Z}/p\mathbb{Z})^m$.

g. Résumer le résultat obtenu.

E15. Soit G un groupe fini d'ordre p^n (p premier), et soit X l'ensemble des parties de G à k éléments ($k \in [\![0, p^n]\!]$).

a. Si $\mathcal{P} \in X$, montrer que $g\mathcal{P} = \{gx \mid x \in \mathcal{P}\}$ est aussi un élément de X.

b. Vérifier que l'application

$$\begin{aligned} G \times X &\longrightarrow X \\ (g, \mathcal{P}) &\longmapsto g\mathcal{P} \end{aligned}$$

définit une action de G sur X.

c. On suppose que $k \in [\![1, p^n - 1]\!]$. Montrer par l'absurde que G agit sans points fixes.

d. En déduire que pour tout $k \in [\![1, p^n - 1]\!]$, on a $\binom{p^n}{k} \equiv 0 \ [p]$.

E16. Soit p un nombre premier et soient $r \geqslant 0$ et $m \geqslant 1$ des entiers. Soit S un groupe d'ordre p^r et soit T un ensemble à m éléments. Soit $X = S \times T$, et soit E l'ensemble des parties de X à p^r éléments.

a. Montrer que l'application

$$\begin{aligned} S \times X &\longrightarrow X \\ \big(s,(s',t)\big) &\longmapsto (ss',t) \end{aligned}$$

définit une action de S sur X.

Pour $\mathcal{P} \in E$, et $s \in S$, on pose $s{\cdot}\mathcal{P} = \{s{\cdot}x \mid x \in \mathcal{P}\}$. On vérifie aisément que l'on définit ainsi une action de S sur E.

b. Déterminer l'ensemble des points fixes de E sous cette action et son cardinal.

c. En déduire la congruence $\binom{p^r m}{p^r} \equiv m \ \ [p]$.

E17. Soit p un nombre premier. Montrer qu'un groupe fini G est d'ordre une puissance de p si, et seulement si, tout élément de G est d'ordre une puissance de p.

E18. Soit G un groupe, et soient H, K deux sous-groupes de G.

a. On suppose que K est d'indice fini dans G. En considérant l'action de H sur l'ensemble G/K par translation à gauche, montrer que $H \cap K$ est d'indice fini dans H, et que l'on a

$$[H : H \cap K] \leqslant [G : K],$$

avec égalité si, et seulement si, on a $G = HK = \{hk \mid h \in H, k \in K\}$.

Indication. On pourra s'intéresser au stabilisateur de $\overline{1}_G$ sous l'action de H.

b. En déduire que $H \cap K$ est d'indice fini dans G si, et seulement si, H et K le sont. Montrer que, dans ce cas, on a

$$\mathrm{ppcm}([G : H], [G : K]) \leqslant [G : H \cap K] \leqslant [G : H][G : K],$$

et que $[G : H \cap K] = [G : H]\times[G : K]$ si, et seulement si, $G = HK$.

c. On suppose que H et K sont d'indices finis, et que $[G : H]$ et $[G : K]$ sont premiers entre eux. Montrer que $[G : H \cap K] = [G : H]\times[G : K]$.

E19. Soit K un corps, soit V un K-espace vectoriel et soit G un groupe. On suppose que le groupe G agit linéairement sur V, c'est-à-dire que

l'action de G sur V vérifie

$$g\cdot(v_1+v_2)=g\cdot v_1+g\cdot v_2, \qquad g\cdot(\lambda v)=\lambda(g\cdot v),$$

pour tous $v_1, v_2, v\in V$, et tout $\lambda\in K$.

a. Montrer que l'ensemble des points fixes V^G est un sous-espace vectoriel de V.

b. Montrer que l'image du morphisme de groupes $\rho : G\longrightarrow \mathfrak{S}(V)$ induit par l'action de G est incluse dans $\mathrm{GL}(V)$.

On note encore $\rho : G\longrightarrow \mathrm{GL}(V)$ le morphisme correspondant. On suppose dorénavant que G est fini et que $\mathrm{car}(K)\nmid |G|$.

c.

(i) Montrer que $p_G=\dfrac{1}{|G|}\displaystyle\sum_{g\in G}\rho(g)$ est un projecteur sur V^G.

(ii) En déduire que si V est de dimension finie, on a

$$\mathrm{tr}(p_G)=\dim_K(V^G)\cdot 1_K.$$

d. Soit E un ensemble fini, sur lequel G agit. Soit $V=\mathscr{F}(E,K)$ l'ensemble des fonctions de E dans K. Pour tout $g\in G$, et tout $f\in V$, on note $g\cdot f$ la fonction définie par

$$(g\cdot f)(x)=f(g^{-1}\cdot x), \qquad \text{pour tout } x\in E.$$

(i) Montrer que cela définit une action linéaire de G sur V.

(ii) Pour tout $x\in E$, on note $e_x\in V$ la fonction définie par $e_x(x)=1$ et $e_x(y)=0$ si $y\neq x$.
Montrer que $(e_x)_{x\in E}$ est une base de V.

(iii) Montrer que pour tout $g\in G$, on a

$$\mathrm{tr}\big(\rho(g)\big)=|\mathrm{Fix}(g)|\cdot 1_K,$$

où $\mathrm{Fix}(g)=\{x\in E \mid g\cdot x=x\}$.

(iv) Montrer que $f\in V^G$ si, et seulement si, f est constante sur les orbites de E sous l'action de G. En déduire que

$$\dim_K(V^G)=|\Omega|,$$

où Ω est l'ensemble des orbites de E sous l'action de G.

(v) En déduire une nouvelle démonstration de la formule de Burnside.

E20. On suppose que X possède au moins deux éléments, et que G agit transitivement sur X. En utilisant la formule de Burnside, montrer qu'il existe un élément $g\in G$ tel que $g\cdot x\neq x$, pour tout $x\in X$.

Démonstration. Soit $f : E \longrightarrow E'$ une bijection de E sur E'. On vérifie facilement que l'application

$$\begin{aligned} u : \mathfrak{S}(E) &\longrightarrow \mathfrak{S}(E') \\ \sigma &\longmapsto f \circ \sigma \circ f^{-1} \end{aligned}$$

est un isomorphisme de groupes. Les détails sont laissés au lecteur. □

1.3. Exemple. Soit E un ensemble à n éléments, et soient $a_1, \ldots, a_n$ les éléments de E. L'application

$$\begin{aligned} f : [\![1, n]\!] &\longrightarrow E \\ i &\longmapsto a_i \end{aligned}$$

est alors bijective, et induit donc un isomorphisme de groupes

$$\begin{aligned} u : \mathfrak{S}_n &\xrightarrow{\sim} \mathfrak{S}(E) \\ \sigma &\longmapsto f \circ \sigma \circ f^{-1}. \end{aligned}$$

Si $\sigma \in \mathfrak{S}_n$, alors on voit facilement que

$$u(\sigma)(a_i) = a_{\sigma(i)}, \quad \text{pour } i \in [\![1, n]\!].$$

Le but de ce chapitre est d'étudier $\mathfrak{S}(E)$ lorsque E est un ensemble fini ayant au moins deux éléments (sinon $\mathfrak{S}(E)$ est le groupe trivial). Rappelons que si E a n éléments, alors $\mathfrak{S}(E)$ est d'ordre $n!$.

Au vu du lemme précédent, on peut se restreindre à l'étude du groupe $\mathfrak{S}_n$, ce que nous ferons parfois.

1.4. Définition. Soit E un ensemble non vide. Le *support* d'une permutation $\sigma \in \mathfrak{S}(E)$ est l'ensemble

$$\mathrm{Supp}(\sigma) = \{a \in E \mid \sigma(a) \neq a\}.$$

Autrement dit, $\mathrm{Supp}(\sigma) = E \setminus \mathrm{Fix}(\sigma)$, où $\mathrm{Fix}(\sigma)$ est l'ensemble des points fixes de σ.

Le support de σ est vide si, et seulement si, σ est l'identité.

Le lemme suivant résume les propriétés du support.

1.5. Lemme. *Soit E un ensemble non vide.*

(1) *Pour tout $\sigma \in \mathfrak{S}(E)$, l'ensemble* $\mathrm{Supp}(\sigma)$ *est stable par σ. On a même*

$$\sigma\big(\mathrm{Supp}(\sigma)\big) = \mathrm{Supp}(\sigma).$$

(2) *Pour tout $\sigma \in \mathfrak{S}(E)$, on a* $\mathrm{Supp}(\sigma) = \mathrm{Supp}(\sigma^{-1})$.

(3) *Pour tout $\sigma \in \mathfrak{S}(E)$, et pour tout $m \in \mathbb{Z}$, on a*

$$\mathrm{Supp}(\sigma^m) \subset \mathrm{Supp}(\sigma).$$

(4) *Pour tous* $\sigma, \sigma' \in \mathfrak{S}(E)$, *on a*

$$\mathrm{Supp}(\sigma\sigma') \subset \mathrm{Supp}(\sigma) \cup \mathrm{Supp}(\sigma').$$

Si de plus σ *et* σ' *sont à supports disjoints, alors*

$$\mathrm{Supp}(\sigma\sigma') = \mathrm{Supp}(\sigma) \cup \mathrm{Supp}(\sigma').$$

Démonstration

(1) Pour tout $a \in E$, on a $\sigma\big(\sigma(a)\big) = \sigma(a) \iff \sigma(a) = a$. On en déduit le résultat voulu par passage au complémentaire.

(2) Pour tout $a \in E$, on a

$$\sigma(a) = a \quad \Longleftrightarrow \quad a = \sigma^{-1}(a).$$

On conclut comme précédemment.

(3) C'est clair par passage au complémentaire, puisque tout élément fixé par σ est fixé par σ^{-1}, donc par σ^m (quelque soit le signe de m).

(4) Soient $\sigma, \sigma' \in \mathfrak{S}(E)$. Si $a \in E$ est fixé par σ et σ', il est fixé par $\sigma\sigma'$, d'où la première inclusion par passage au complémentaire.

Supposons maintenant que σ et σ' soient à supports disjoints.

Soit $a \in \mathrm{Supp}(\sigma) \cup \mathrm{Supp}(\sigma')$.

Si $a \in \mathrm{Supp}(\sigma)$, on a $\sigma(a) \neq a$. Par hypothèse, a est alors fixé par σ', et donc

$$\sigma\sigma'(a) = \sigma(a) \neq a.$$

Ainsi, $a \in \mathrm{Supp}(\sigma\sigma')$.

Si $a \in \mathrm{Supp}(\sigma')$, alors $\sigma'(a) \neq a$. De plus, $\sigma'(a) \in \mathrm{Supp}(\sigma')$ par (1), et $\sigma'(a)$ est alors fixé par σ par hypothèse. Ainsi, on a

$$\sigma\big(\sigma'(a)\big) = \sigma'(a) \neq a,$$

ce qui prouve que $a \in \mathrm{Supp}(\sigma\sigma')$, et établit l'égalité voulue. □

Le groupe $\mathfrak{S}(E)$ n'est pas abélien dès que E possède au moins trois éléments. En revanche, on a le résultat suivant.

1.6. Lemme. *Deux permutations à supports disjoints commutent.*

Démonstration. Soient σ, $\sigma' \in \mathfrak{S}(E)$ deux permutations à supports disjoints et soit $a \in E$. Si a est fixé par σ et par σ', alors on a

$$\sigma\big(\sigma'(a)\big) = \sigma(a) = a = \sigma'(a) = \sigma'\big(\sigma(a)\big).$$

Si $a \in \mathrm{Supp}(\sigma)$, alors $\sigma(a) \in \mathrm{Supp}(\sigma)$ d'après le lemme 1.5 (1). Par hypothèse, $\sigma(a)$ n'est pas dans le support de σ', et donc $\sigma'\big(\sigma(a)\big) = \sigma(a)$.

Mais, puisque $a \in \mathrm{Supp}(\sigma)$, a n'est pas dans le support de σ', et l'on a donc

$$\sigma'\big(\sigma(a)\big) = \sigma(a) = \sigma\big(\sigma'(a)\big).$$

On raisonne de la même manière lorsque $a \in \mathrm{Supp}(\sigma')$. □

2. Décomposition en produit de cycles

Dans tout ce qui suit, E est un ensemble fini à n éléments, où $n \geqslant 2$.

Soit $\sigma \in \mathfrak{S}(E)$. On rappelle que l'action naturelle de $\mathfrak{S}(E)$ sur E se restreint en une action du sous-groupe $\langle\sigma\rangle$ sur E, et qu'une orbite de E sous cette action s'appelle une *σ-orbite* de E. Si $a \in E$, on notera $\mathrm{Orb}_\sigma(a)$ la σ-orbite correspondante. Par définition, on a donc

$$\mathrm{Orb}_\sigma(a) = \{\sigma^m(a) \mid m \in \mathbb{Z}\}.$$

Il est clair que $\mathrm{Orb}_\sigma(a)$ est réduite à un singleton si, et seulement si, a est un point fixe de σ. Les différentes σ-orbites de E forment alors une partition de E, et la réunion des σ-orbites non réduites à un singleton est égale au support de σ. Puisque $\mathfrak{S}(E)$ est fini, tout sous-groupe est d'indice fini, et le lemme VII-1.14 nous donne immédiatement le résultat suivant.

2.1. Lemme. *Soit $\sigma \in \mathfrak{S}(E)$ une permutation, soit ω une σ-orbite à p éléments et soit $a \in \omega$. Alors, p est le plus petit entier strictement positif tel que $\sigma^p(a) = a$, et l'on a*

$$\omega = \{a,\ \sigma(a),\ \ldots,\ \sigma^{p-1}(a)\}.$$

On définit maintenant la notion de cycle.

2.2. Définition. On dit qu'une permutation $\sigma \in \mathfrak{S}(E)$ est un *cycle* s'il n'existe qu'une seule σ-orbite non réduite à un singleton (nécessairement égale à son support). Si $p \geqslant 2$ est le nombre d'éléments de $\mathrm{Supp}(\sigma)$, on dit que σ est un p-cycle.

Un 2-cycle est appelé une *transposition*. Autrement dit, une transposition de $\mathfrak{S}(E)$ a $n-2$ points fixes et échange deux éléments.

2.3. Exemple. Soit $p \geqslant 2$, et soient $a_1, \ldots, a_p \in E$ p éléments distincts. La permutation $\sigma \in \mathfrak{S}(E)$ définie par

$$\sigma(a_1) = a_2,\ \ldots,\ \sigma(a_{p-1}) = a_p,\ \sigma(a_p) = a_1, \quad \sigma(a) = a,\ \text{ si } a \neq a_i,\ i \in [\![1, p]\!]$$

est un p-cycle de support $\{a_1, \ldots, a_p\}$.

Pour le voir, il faut montrer qu'il y a une unique σ-orbite non réduite à un élément. Soit $a \in E$. Si $a \neq a_i, i \in [\![1,p]\!]$, alors a est un point fixe sous σ, et sa σ-orbite est réduite à un élément. Pour conclure, il suffit de voir que les éléments $a_1, \ldots, a_p \in E$ forment une seule σ-orbite. Mais ces éléments sont tous dans la σ-orbite de a_1 puisque l'on a

$$\sigma^{i-1}(a_1) = a_i, \quad \text{pour tout } i \in [\![1,p]\!],$$

d'où le résultat.

Notation. Le p-cycle de l'exemple précédent sera noté $(a_1 \cdots a_p)$.

2.4. Exemple. Soit $E = [\![1,4]\!]$. Alors, $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}$ est un 3-cycle. C'est en fait le 3-cycle (1 3 2). Remarquons que l'on a aussi

$$\sigma = (1\ 3\ 2) = (2\ 1\ 3) = (3\ 2\ 1).$$

En général, on décide de démarrer par le plus petit élément dans le support. De même, $\tau = (2\ 4) = (4\ 2)$ est une transposition, définie par

$$\tau(1) = 1, \quad \tau(2) = 4, \quad \tau(3) = 3, \quad \tau(4) = 2.$$

On va maintenant montrer que toute permutation de $\mathfrak{S}(E)$ se décompose en produit de cycles à supports disjoints. L'idée est que pour connaître une permutation, il suffit de connaître son action sur chaque orbite non réduite à un élément. La première étape est d'associer à chaque orbite ω un cycle de support ω.

Soit $\sigma \in \mathfrak{S}(E)$. Soit Ω^* l'ensemble des σ-orbites non réduites à un élément. Si $\omega \in \Omega^*$, on pose

$$\sigma_\omega(a) = \sigma(a), \quad \text{si } a \in \omega \quad \text{et} \quad \sigma_\omega(a) = a, \quad \text{si } a \notin \omega.$$

2.5. Lemme. *Soit $\sigma \in \mathfrak{S}(E)$ et soit $\omega \in \Omega^*$ une orbite non réduite à un élément. Alors, σ_ω est un cycle de support ω, et pour tout $a \in \omega$, on a*

$$\sigma_\omega = \big(a\ \sigma(a)\ \ldots\ \sigma^{p-1}(a)\big),$$

où p est le nombre d'éléments de ω.

De plus, tout p-cycle de $\mathfrak{S}(E)$ est de cette forme.

Démonstration. Soit $\omega \in \Omega^*$ une σ-orbite de cardinal $p > 1$. Soit $a \in \omega$. Par le lemme 2.1, p est le plus petit entier > 0 tel que $\sigma^p(a) = a$, et l'on a

$$\omega = \{a,\ \sigma(a),\ \ldots,\ \sigma^{p-1}(a)\}.$$

On va montrer l'égalité

$$\sigma_\omega = \big(a\ \sigma(a)\ \ldots\ \sigma^{p-1}(a)\big).$$

Soit $a' \in E$. Si $a' \in E \setminus \{a, \sigma(a), \ldots, \sigma^{p-1}(a)\}$, alors $a' \notin \omega$, et $\sigma_\omega(a') = a'$ par définition. De plus, on a

$$\sigma_\omega\big(\sigma^i(a)\big) = \sigma^{i+1}(a), \quad \text{pour tout } i \in [\![0, p-1]\!]$$

puisque $\sigma^i(a) \in \omega$. Puisque $\sigma^p(a) = a$, cela montre l'égalité voulue. Le fait que le support de σ_ω soit égal à ω provient de l'égalité précédente et de l'exemple VIII-2.3.

Si maintenant $\sigma \in \mathfrak{S}(E)$ est un p-cycle, posons $\omega = \mathrm{Supp}(\sigma)$. Montrons que $\sigma_\omega = \sigma$. Si $a \in \omega$, on a $\sigma_\omega(a) = \sigma(a)$ par définition de σ_ω, et si $a \notin \omega$, on a $\sigma_\omega(a) = a = \sigma(a)$, la dernière égalité provenant du fait que $E \setminus \omega$ est l'ensemble des points fixes de σ. □

2.6. Remarque. Le lemme précédent montre que tout p-cycle de $\mathfrak{S}(E)$ est de la forme $(a_1 \ldots a_p)$, où $a_1, \ldots, a_p$ sont p éléments distincts de E.

On a alors le théorème suivant.

2.7. Théorème. *Soit $\sigma \in \mathfrak{S}(E)$. Alors, σ se décompose en produit de cycles à supports disjoints, et cette décomposition est unique à l'ordre des facteurs près. Cette décomposition est donnée par*

$$\sigma = \prod_{\omega \in \Omega^*} \sigma_\omega,$$

où Ω^* est l'ensemble des σ-orbites non réduites à un singleton.

Démonstration. Soit $\sigma \in \mathfrak{S}(E)$. Montrons l'égalité

$$\sigma = \prod_{\omega \in \Omega^*} \sigma_\omega.$$

Remarquons tout d'abord que les σ_ω sont des cycles dont les supports sont disjoints deux à deux. En effet, $\mathrm{Supp}(\sigma_\omega) = \omega$, et deux σ-orbites distinctes sont disjointes.

Soit $a \in E$. Si a est un point fixe de σ, alors $a \notin \omega$ pour tout $\omega \in \Omega^*$ (sinon ω serait réduite à un élément). Ainsi, $\sigma_\omega(a) = a$, pour tout $\omega \in \Omega^*$, et donc

$$\prod_{\omega \in \Omega^*} \sigma_\omega(a) = a = \sigma(a).$$

Si a est dans le support de σ, alors $\mathrm{Orb}_\sigma(a) \in \Omega^*$ et donc il existe $\omega_0 \in \Omega^*$ tel que $\mathrm{Orb}_\sigma(a) = \omega_0$. On a donc $a \in \omega_0$, et $a \notin \omega$ pour tout $\omega \in \Omega^* \setminus \{\omega_0\}$, car deux σ-orbites distinctes sont disjointes. Par conséquent, $\sigma_\omega(a) = a$, pour tout $\omega \neq \omega_0$. Comme les supports des cycles σ_ω sont disjoints, ils commutent.

Ainsi, on a

$$\prod_{\omega\in\Omega^*}\sigma_\omega(a)=\sigma_{\omega_0}\Big(\Big(\prod_{\omega\in\Omega^*\setminus\{\omega_0\}}\sigma_\omega\Big)(a)\Big)=\sigma_{\omega_0}(a)=\sigma(a),$$

la dernière égalité provenant de la définition de σ_{ω_0}.

Montrons maintenant l'unicité de la décomposition. Supposons que l'on ait une décomposition

$$\sigma=\sigma_1\cdots\sigma_r,$$

où les σ_i sont des cycles à supports disjoints. On note E_i le support du cycle σ_i. On va montrer que E_i est une σ-orbite non réduite à un élément.

Soit $a\in E_i$. Comme les supports des σ_i sont disjoints, on a $\sigma_j(a)=a$ dès que $j\neq i$. De plus, puisque les σ_j commutent, on a

$$\sigma(a)=\Big(\prod_j\sigma_j\Big)(a)=\sigma_i\Big(\Big(\prod_{j\neq i}\sigma_j\Big)(a)\Big)=\sigma_i(a).$$

On a donc $\sigma_i(a)=\sigma(a)$ si $a\in E_i$. Puisque $E_i=\mathrm{Supp}(\sigma_i)$ est stable par σ_i par le lemme 1.5 (1), pour tout $a\in E_i$, on a donc $\sigma(a)=\sigma_i(a)\in E_i$. Par récurrence, on a $\sigma_i^k(a)=\sigma^k(a), k\geqslant 0$.

Comme σ_i est un cycle de support E_i, alors E_i forme une σ_i-orbite non réduite à un élément. Le lemme 2.1 appliqué à σ_i montre alors que le support est $E_i=\{a,\sigma_i(a),\ldots,\sigma_i^{p_i-1}(a)\}$, où p_i est le cardinal de E_i. Par ce qui précède, on obtient donc

$$E_i=\{a,\ \sigma(a),\ \ldots,\ \sigma^{p_i-1}(a)\},$$

ce qui montre que les éléments de E_i forment une σ-orbite (à savoir celle de a) non réduite à un élément.

Remarquons alors que, puisque le support de σ est la réunion des E_i, on obtient bien toutes les orbites non réduites à un élément. Cela prouve l'unicité de la décomposition, et achève la démonstration. □

2.8. Remarque. Les deux résultats précédents fournissent l'égalité suivante : soit $\sigma\in\mathfrak{S}(E)$, et soient $\omega_1,\ldots,\omega_r$ les différentes σ-orbites non réduites à un élément. Soit $a_i\in\omega_i$, et soit p_i le cardinal de ω_i.

Alors, on a

$$\sigma=\big(a_1\ \sigma(a_1)\ \ldots\ \sigma^{p_1-1}(a_1)\big)\cdots\big(a_r\ \sigma(a_r)\ \ldots\ \sigma^{p_r-1}(a_r)\big).$$

Tout cela donne une méthode systématique pour trouver la décomposition de σ. Pour fixer les idées, supposons que $E=[\![1,n]\!]$. Si $\sigma\in\mathfrak{S}_n$, on regarde l'image de 1 par les puissances successives de σ jusqu'à reboucler sur 1. Ensuite, on prend un autre entier qui n'est pas encore atteint et l'on recommence le procédé, jusqu'à épuiser tous les entiers de 1 à n.

2.9. Exemple. Soit $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 6 & 2 & 4 & 1 & 3 \end{pmatrix} \in \mathfrak{S}_6$.

Alors, on a

$$\sigma(1) = 5 \quad \text{et} \quad \sigma(5) = 1.$$

On obtient un premier cycle de longueur 2, à savoir (1 5).

On choisit maintenant un entier distinct de 1 et 5, disons 2. On a

$$\sigma(2) = 6, \quad \sigma(6) = 3 \quad \text{et} \quad \sigma(3) = 2,$$

et l'on obtient le cycle (2 6 3).

Le seul entier non encore utilisé est 4. Or, on a $\sigma(4) = 4$. Ainsi, 4 n'est pas dans le support de σ. On a donc

$$\sigma = (1\ 5)(2\ 6\ 3).$$

Le théorème 2.7 entraîne immédiatement le résultat suivant.

2.10. Corollaire. *Soit E un ensemble fini. Le groupe $\mathfrak{S}(E)$ est engendré par les cycles.*

Au vu de ce résultat, on donne souvent un élément de $\mathfrak{S}(E)$ sous forme de produits de cycles, pas nécessairement à supports disjoints. Il faut vite s'habituer à calculer l'image par σ à partir d'une telle décomposition et à en déduire une décomposition en produits de cycles à supports disjoints.

2.11. Exemple. Soit $\sigma \in \mathfrak{S}_8$, donné par

$$\sigma = (1\ 2\ 3\ 5)(3\ 7)(7\ 4\ 8).$$

Attention ! On compose des applications, donc on calcule à partir de la droite.

$$\begin{array}{cccccccc} & (7\ 4\ 8) & & (3\ 7) & & (1\ 2\ 3\ 5) & & \\ \underline{1} & \longrightarrow & 1 & \longrightarrow & 1 & \longrightarrow & \underline{2}, \\ \underline{2} & \longrightarrow & 2 & \longrightarrow & 2 & \longrightarrow & \underline{3}, \\ \underline{3} & \longrightarrow & 3 & \longrightarrow & 7 & \longrightarrow & \underline{7}, \\ \underline{7} & \longrightarrow & 4 & \longrightarrow & 4 & \longrightarrow & \underline{4}, \\ \underline{4} & \longrightarrow & 8 & \longrightarrow & 8 & \longrightarrow & \underline{8}, \\ \underline{8} & \longrightarrow & 7 & \longrightarrow & 3 & \longrightarrow & \underline{5}, \\ \underline{5} & \longrightarrow & 5 & \longrightarrow & 5 & \longrightarrow & \underline{1}. \end{array}$$

On obtient un premier cycle (1 2 3 7 4 8 5). Il reste donc à regarder l'image de 6, qui est fixe. On a donc $\sigma = (1\ 2\ 3\ 7\ 4\ 8\ 5)$.

2.12. Corollaire. *Le groupe $\mathfrak{S}(E)$ est engendré par les transpositions.*

Démonstration. Puisque toute permutation est produit de cycles, il suffit de montrer qu'un cycle est un produit de transpositions. Or, un simple

calcul montre que l'on a

$$(a_1 \cdots a_p) = (a_1\ a_2)(a_2\ a_3)\cdots(a_{p-2}\ a_{p-1})(a_{p-1}\ a_p),$$

d'où le résultat. □

2.13. Remarque. On peut retrouver ce fait de manière plus directe, en utilisant le fait que les transpositions engendrent $\mathfrak{S}_n$ (cf. exemple VI-3.8).

En effet, soit $a_1, \ldots, a_n$ les éléments de E, et soit

$$\begin{aligned} f : [\![1, n]\!] &\longrightarrow E \\ i &\longmapsto a_i. \end{aligned}$$

On vérifie l'isomorphisme de groupes

$$\begin{aligned} \mathfrak{S}_n &\longrightarrow \mathfrak{S}(E) \\ \sigma &\longmapsto f \circ \sigma \circ f^{-1} \end{aligned}$$

envoie la transposition $(i\ j) \in \mathfrak{S}_n$ sur la transposition $(a_i\ a_j) \in \mathfrak{S}(E)$. Comme les transpositions engendrent $\mathfrak{S}_n$, on en déduit aisément le résultat voulu.

Les détails sont laissés au lecteur.

On continue ce paragraphe par un résultat qui calcule l'ordre d'une permutation quelconque. On commence par démontrer deux lemmes.

2.14. Lemme. *Un p-cycle est d'ordre p.*

Démonstration. Soit $\sigma = (a_1 \ldots a_p)$. Pour $j \in [\![1, p-1]\!]$, on a

$$\sigma^j(a_1) = a_{j+1} \neq a_1.$$

Donc $\sigma, \sigma^2, \ldots, \sigma^{p-1}$ ne sont pas égaux à Id_E. En revanche, pour $k \in [\![1, p]\!]$, on a

$$\sigma^p(a_k) = \sigma^p\big(\sigma^{k-1}(a_1)\big) = \sigma^{k-1}\big(\sigma^p(a_1)\big) = \sigma^{k-1}(a_1) = a_k,$$

d'où $\sigma^p = \mathrm{Id}_E$. Ainsi, $o(\sigma) = p$ par le théorème VI-4.13. □

2.15. Lemme. *Soient $\sigma_1, \ldots, \sigma_r \in \mathfrak{S}(E)$ des permutations à supports deux à deux disjoints. Alors, on a*

$$o(\sigma_1 \cdots \sigma_r) = \mathrm{ppcm}\big(o(\sigma_1), \ldots, o(\sigma_r)\big).$$

Démonstration. Procédons par récurrence sur r. Si $r = 1$, le résultat est vrai. Montrons-le également pour $r = 2$. Soient $\sigma_1, \sigma_2 \in \mathfrak{S}(E)$ deux permutations à supports disjoints. Par le lemme 1.6, σ_1 et σ_2 commutent. Par ailleurs, on a

$$\langle \sigma_1 \rangle \cap \langle \sigma_2 \rangle = \{\mathrm{Id}_E\}.$$

En effet, soit $\sigma \in \langle \sigma_1 \rangle \cap \langle \sigma_2 \rangle$. En utilisant le lemme 1.5 (3), on obtient que

$$\mathrm{Supp}(\sigma) \subset \mathrm{Supp}(\sigma_1) \cap \mathrm{Supp}(\sigma_2).$$

L'hypothèse implique alors que σ est de support vide, c'est-à-dire $\sigma = \mathrm{Id}_E$. D'après la proposition VI-4.17, on en déduit l'égalité

$$o(\sigma_1\sigma_2) = \mathrm{ppcm}(o(\sigma_1), o(\sigma_2)).$$

Supposons maintenant que l'égalité soit vraie au rang $r \geqslant 1$, et montrons-le au rang $r+1$. Si $\sigma_1, \ldots, \sigma_{r+1} \in \mathfrak{S}(E)$ sont à supports deux à deux disjoints, alors $\sigma_1 \cdots \sigma_r$ et σ_{r+1} sont aussi à supports disjoints.

En effet, une utilisation répétée du lemme 1.5 (4) montre que l'on a[1]

$$\mathrm{Supp}(\sigma_1 \cdots \sigma_r) \subset \mathrm{Supp}(\sigma_1) \cup \cdots \cup \mathrm{Supp}(\sigma_r).$$

L'hypothèse implique que l'ensemble de droite n'intersecte pas $\mathrm{Supp}(\sigma_{r+1})$, d'où le fait annoncé. Mais alors, le cas $r = 2$ nous donne l'égalité

$$o(\sigma_1 \cdots \sigma_{r+1}) = o\big((\sigma_1 \cdots \sigma_r)\sigma_{r+1}\big) = \mathrm{ppcm}\big(o(\sigma_1 \cdots \sigma_r), o(\sigma_{r+1})\big).$$

Il suffit alors d'appliquer l'hypothèse de récurrence et les propriétés d'associativité du ppcm pour conclure. □

Les deux lemmes précédents entraînent alors immédiatement le résultat suivant.

2.16. Théorème. *L'ordre d'une permutation est le ppcm des longueurs des cycles à supports disjoints qui la composent.*

Ainsi, pour calculer l'ordre de $\sigma \in \mathfrak{S}(E)$, il suffit donc de la décomposer en produit de cycles à supports disjoints, et d'appliquer le théorème précédent, ce qui est beaucoup plus rapide que de chercher le plus petit entier $p \geqslant 1$ tel que $\sigma^p = \mathrm{Id}_E$.

On termine ce paragraphe en déterminant les classes de conjugaison dans le groupe $\mathfrak{S}(E)$.

1. Il y a même égalité, mais nous n'en aurons pas besoin.

2.17. Lemme. *Soit $\tau \in \mathfrak{S}(E)$. Pour tout p-cycle $(a_1 \dots a_p)$, on a*

$$\tau(a_1 \dots a_p)\tau^{-1} = \big(\tau(a_1) \dots \tau(a_p)\big).$$

En particulier, le conjugué d'un cycle est encore un cycle de même longueur.

Démonstration. Soit $\tau' = \tau(a_1 \dots a_p)\tau^{-1}$. Alors, on a

$$\tau'\big(\tau(a_i)\big) = \tau\big((a_1 \dots a_p)(a_i)\big) = \tau(a_{i+1}), \quad \text{si } i \in [\![1, p-1]\!],$$

et

$$\tau'\big(\tau(a_p)\big) = \tau\big((a_1 \dots a_p)(a_p)\big) = \tau(a_1), \quad \text{si } i = p.$$

Soit maintenant $a \in E \setminus \{\tau(a_1), \dots, \tau(a_p)\}$. Puisque τ est bijective, donc injective, on en déduit que $\tau^{-1}(a) \neq a_i$, pour tout $i \in [\![1, p]\!]$. En particulier, le p-cycle $(a_1 \dots a_p)$ fixe $\tau^{-1}(a)$, et l'on a donc

$$\tau'(a) = \tau\big(\tau^{-1}(a)\big) = a.$$

Cela achève la démonstration. □

2.18. Théorème. *Deux permutations sont conjuguées dans $\mathfrak{S}(E)$ si, et seulement si, les listes (avec répétition) des longueurs des cycles à supports disjoints qui les composent sont les mêmes à l'ordre près.*

Démonstration. Soit $\sigma = \sigma_1\sigma_2 \cdots \sigma_r$, où les σ_i sont des cycles à supports disjoints. Soit $\sigma' \in \mathfrak{S}(E)$. Supposons que σ et σ' soient conjuguées. Alors, il existe $\tau \in \mathfrak{S}(E)$ tel que $\sigma' = \tau\sigma\tau^{-1}$. On a

$$\tau\sigma\tau^{-1} = (\tau\sigma_1\tau^{-1})(\tau\sigma_2\tau^{-1}) \cdots (\tau\sigma_r\tau^{-1}).$$

D'après le lemme précédent, $\sigma'_i = \tau\sigma_i\tau^{-1}$ est un cycle de même longueur que σ_i. De plus, toujours par le même lemme, le support de σ' est l'image du support de σ_i par τ. Puisque τ est bijective et que les supports des σ_i sont disjoints, il en est de même pour les supports des σ'_i. Ainsi, $\sigma' = \sigma'_1 \cdots \sigma'_r$ est la décomposition de σ' en produits de cycles à supports disjoints, et les listes des longueurs des cycles à supports disjoints qui composent σ et σ' sont les mêmes.

Réciproquement, soient σ, $\sigma' \in \mathfrak{S}(E)$. Considérons alors leurs décompositions en produits de cycles à supports disjoints

$$\sigma = \sigma_1 \cdots \sigma_r, \quad \sigma' = \sigma'_1 \cdots \sigma'_r,$$

et supposons que pour tout $i \in [\![1, r]\!]$, les cycles σ_i et σ'_i soient de même longueur, quitte à renuméroter.

Écrivons

$$\sigma_i = \big(a_1^{(i)} \dots a_{p_i}^{(i)}\big), \quad \sigma'_i = \big(b_1^{(i)} \dots b_{p_i}^{(i)}\big).$$

Puisque

$$\mathrm{Supp}(\sigma) = \bigcup_{i=1}^{r} \mathrm{Supp}(\sigma_i) \quad \text{et} \quad \mathrm{Supp}(\sigma') = \bigcup_{i=1}^{r} \mathrm{Supp}(\sigma'_i),$$

par le lemme 1.5 (4), et que ces deux réunions sont disjointes, l'hypothèse entraîne que $\mathrm{Supp}(\sigma)$ et $\mathrm{Supp}(\sigma')$ ont même cardinal. Il en est donc de même des ensembles $X = E \setminus \mathrm{Supp}(\sigma)$ et $X' = E \setminus \mathrm{Supp}(\sigma')$. Choisissons une bijection $f : X \longrightarrow X'$ arbitraire, et soit $\tau : E \longrightarrow E$ définie par

$$\tau(a_j^{(i)}) = b_j^{(i)} \quad \text{et} \quad \tau(a) = f(a), \ \text{ si } a \neq a_j^{(i)}.$$

Remarquons que cela définit un élément de $\mathfrak{S}(E)$, par construction. Par définition de τ, et par le lemme précédent, on a

$$\tau\sigma_i\tau^{-1} = \sigma'_i, \quad \text{pour tout } i \in [\![1, r]\!].$$

On a alors

$$\tau\sigma\tau^{-1} = (\tau\sigma_1\tau^{-1})(\tau\sigma_2\tau^{-1})\cdots(\tau\sigma_r\tau^{-1}) = \sigma'_1\cdots\sigma'_r = \sigma'.$$

Cela achève la démonstration. □

2.19. Exemple. Les éléments $\sigma = (1\ 3\ 5)(2\ 4)$ et $\sigma' = (1\ 2\ 3)(6\ 7)$ sont conjugués dans $\mathfrak{S}_7$. Pour trouver un élément τ qui convient, on applique la méthode de la démonstration. On a

$$\begin{aligned} \sigma &= (1\ 3\ 5)(2\ 4) \\ \sigma' &= (1\ 2\ 3)(6\ 7). \end{aligned}$$

Les point fixes de σ sont 6 et 7, et ceux de σ' sont 4 et 5. Ainsi, on peut prendre

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 1 & 6 & 2 & 7 & 3 & 4 & 5 \end{pmatrix} = (2\ 6\ 4\ 7\ 5\ 3),$$

car τ induit bien une bijection de $\{6, 7\}$ (l'ensemble des points fixes de σ) sur $\{4, 5\}$ (l'ensemble des points fixes de σ').

Si l'on préfère, pour faciliter les calculs, on peut rajouter les points fixes dans l'écriture en étendant la notation $(a_1 \ldots a_p)$ à $p = 1$. Attention néanmoins, (a) n'est pas un cycle!! C'est en fait Id_E. Ainsi, on pourra écrire

$$\begin{aligned} \sigma &= (1\ 3\ 5)(2\ 4)(6)(7) \\ \sigma' &= (1\ 2\ 3)(6\ 7)(4)(5), \end{aligned}$$

ce qui permet de « lire » plus facilement comment doit agir τ sur les points fixes.

Nous allons maintenant donner un résultat qui permet de lister toutes les classes de conjugaison de $\mathfrak{S}(E)$. Donnons tout d'abord quelques définitions.

2.20. Définition. Soit $n \geqslant 1$ un entier. Une *partition* de n est une suite d'entiers $\mathbf{p} = (p_k)_{k\geqslant 1}$ décroissante, nulle à partir d'un certain rang, et telle que

$$\sum_{k\geqslant 1} p_k = n.$$

Par abus de notation, on confondra souvent $\mathbf{p}$ avec la suite finie décroissante de ses éléments non nuls.

Notation. On note $\mathscr{P}(n)$ l'ensemble des partitions de l'entier n.

2.21. Exemple. Déterminons toutes les partitions de 4. Puisque

$$4 = 3+1 = 2+2 = 2+1+1 = 1+1+1+1,$$

on obtient cinq partitions, à savoir

$$(4),\quad (3,1),\quad (2,2),\quad (2,1,1)\quad \text{et}\quad (1,1,1,1).$$

2.22. Définition. Si $\sigma \in \mathfrak{S}(E)$, le *type* de σ est la partition de n dont les éléments non nuls sont les cardinaux des diverses σ-orbites, rangés par ordre décroissant. On la note $\mathbf{p}_\sigma$.

2.23. Exemple. Si $\sigma = (1\ 3)(4\ 7\ 6\ 5) \in \mathfrak{S}_9$, on a $\mathbf{p}_\sigma = (4,2,1,1,1)$.

2.24. Remarque. Le théorème 2.18 se retraduit de la manière suivante : deux permutations de $\mathfrak{S}(E)$ sont conjuguées si, et seulement si, elles sont de même type.

En effet, connaître les longueurs des p-cycles à supports disjoints composant σ revient à connaître les cardinaux des σ-orbites non réduites à un singleton. Or, le nombre de σ-orbites réduites à un élément, c'est le nombre de points fixes, c'est-à-dire le nombre d'éléments qui ne sont pas dans le support. Mais, le nombre d'éléments du support est la somme des cardinaux des σ-orbites non réduites à un singleton. Bref, connaître les longueurs des p-cycles à supports disjoints composant σ équivaut à connaître les cardinaux de toutes les σ-orbites, c'est-à-dire connaître son type.

Soit $E = \{a_1, \ldots, a_n\}$ une énumération des éléments de E.

Si $\mathbf{p} = (p_1, \ldots, p_r)$ est une partition de n, on note $\sigma_\mathbf{p} \in \mathfrak{S}(E)$ la permutation

$$\sigma_\mathbf{p} = (a_1 \ldots a_{p_1})(a_{(p_1+1)} \ldots a_{(p_1+p_2)}) \cdots (a_{(p_1+\cdots+p_{r-1}+1)} \ldots a_n).$$

Par convention, si $a \in E$, la notation (a) désignera Id_E. Par construction, la permutation $\sigma_\mathbf{p}$ est alors de type $\mathbf{p}$.

Si $\sigma \in \mathfrak{S}(E)$, on note $\mathrm{Conj}_{\mathfrak{S}(E)}(\sigma)$ la classe de conjugaison de σ. On a alors le corollaire suivant.

2.25. Corollaire. *Soit E un ensemble à n éléments, et soit $\mathscr{C}_E$ l'ensemble des classes de conjugaison de $\mathfrak{S}(E)$. Alors, l'application*

$$\begin{aligned} \Psi : \mathscr{P}(n) &\longrightarrow \mathscr{C}_E \\ \mathbf{p} &\longmapsto \mathrm{Conj}_{\mathfrak{S}(E)}(\sigma_{\mathbf{p}}) \end{aligned}$$

est une bijection de l'ensemble des partitions de n sur l'ensemble des classes de conjugaison de $\mathfrak{S}(E)$.

Autrement dit, toute permutation de $\mathfrak{S}(E)$ est conjuguée à une unique permutation de la forme $\sigma_{\mathbf{p}}$, où $\mathbf{p}$ est une partition de n.

Démonstration. Montrons que Ψ est injective. Soient $\mathbf{p}$ et $\mathbf{p}'$ deux partitions telles que

$$\Psi(\mathbf{p}) = \Psi(\mathbf{p}').$$

Cela signifie donc que les permutations $\sigma_{\mathbf{p}}$ et $\sigma_{\mathbf{p}'}$ sont conjuguées. Or, deux permutations conjuguées sont de même type. Comme $\sigma_{\mathbf{p}}$ et $\sigma_{\mathbf{p}'}$ sont de type $\mathbf{p}$ et $\mathbf{p}'$ respectivement, on a $\mathbf{p} = \mathbf{p}'$, d'où l'injectivité.

Montrons maintenant que Ψ est surjective. Soit $\sigma \in \mathfrak{S}(E)$, et soit $\mathbf{p}_\sigma$ son type. Alors, $\sigma_{\mathbf{p}_\sigma}$ est de type $\mathbf{p}_\sigma$. Ainsi, $\sigma_{\mathbf{p}_\sigma}$ et σ sont de même type, donc conjuguées par la remarque 2.24. On a alors

$$\mathrm{Conj}_{\mathfrak{S}(E)}(\sigma) = \mathrm{Conj}_{\mathfrak{S}(E)}(\sigma_{\mathbf{p}_\sigma}) = \Psi(\mathbf{p}_\sigma),$$

d'où la surjectivité. Cela achève la démonstration. □

Pour trouver toutes les classes de conjugaison de $\mathfrak{S}(E)$, il suffit donc de décomposer n en somme d'entiers strictement positifs décroissants de toutes les façons possibles.

2.26. Exemple. D'après l'exemple 2.21, $\mathfrak{S}_4$ possède cinq classes de conjugaison distinctes, représentées par

$$(1\ 2\ 3\ 4), \quad (1\ 2\ 3), \quad (1\ 2)(3\ 4), \quad (1\ 2), \quad \mathrm{Id}.$$

3. Systèmes de générateurs

Dans ce paragraphe, on donne des systèmes variés de générateurs de $\mathfrak{S}_n$.

3.1. Corollaire. *Le groupe $\mathfrak{S}_n$ est engendré par chacune des familles suivantes :*

(1) *les transpositions* $(1\ i)$, $i \in [\![2, n]\!]$;

(2) *les transpositions* $(i\ i+1)$, $i \in [\![1, n-1]\!]$.

Démonstration. Puisque toute permutation est produit de transpositions, il reste à vérifier que toute transposition est produit de transpositions de

la forme $(1\ i)$ (resp. de la forme $(i\ i+1)$). Or, pour tout $i<j$, on a

$$(i\ j)=(1\ i)(1\ j)(1\ i).$$

L'autre point est à peine plus compliqué.

En utilisant le lemme 2.17, on voit que l'on a

$$(i\ i+1\dots j-1)(j-1\ j)(i\ i+1\dots j-1)^{-1}=(i\ j).$$

Mais comme $(i\ i+1\dots j-1)=(i\ i+1)\cdots(j-2\ j-1)$, on obtient le résultat. □

3.2. Proposition. *Le groupe $\mathfrak{S}_n$ est engendré par $(1\ 2)$ et $(1\ 2\dots n)$.*

Démonstration. Pour $i\in[\![1,n-1]\!]$, la permutation $(1\ 2\dots n)^{i-1}$ envoie 1 sur i et 2 sur $i+1$. Par le lemme 2.17, on a

$$(i\ i+1)=(1\ 2\dots n)^{i-1}(1\ 2)(1\ 2\dots n)^{1-i}.$$

Ainsi, $(i\ i+1)$ est dans le sous-groupe engendré par $(1\ 2)$ et $(1\ 2\dots n)$. On utilise alors la dernière partie du corollaire précédent pour conclure. □

3.3. Remarque. Ce résultat devient faux en général si l'on remplace $(1\ 2)$ et $(1\ 2\dots n)$ par une transposition et un n-cycle arbitraires. Le lecteur pourra chercher un contre-exemple lorsque $n=4$.

4. Signature, groupe alterné

Soit E un ensemble fini à n éléments, $n\geqslant 2$. Nous allons ici définir la signature d'une permutation de E ainsi que le groupe alterné sur E, pour un ensemble fini E quelconque, et présenter un ou deux systèmes de générateurs de ce groupe.

4.1. Théorème. *Soit E un ensemble fini, avec $|E|\geqslant 2$. Alors, il existe un unique morphisme $\varepsilon_E:\mathfrak{S}(E)\longrightarrow\mathbb{C}^\times$ non trivial. Si $\sigma\in\mathfrak{S}(E)$ s'écrit comme produit de s transpositions, alors on a*

$$\varepsilon_E(\sigma)=(-1)^s.$$

Démonstration. Soit $\varepsilon:\mathfrak{S}(E)\longrightarrow\mathbb{C}^\times$ un morphisme de groupes non trivial. Remarquons que ε est constant sur les classes de conjugaison de $\mathfrak{S}(E)$. En effet, si $\sigma'=\tau\sigma\tau^{-1}$, on a

$$\varepsilon(\sigma')=\varepsilon(\tau)\varepsilon(\sigma)\varepsilon(\tau)^{-1}=\varepsilon(\sigma)\varepsilon(\tau)\varepsilon(\tau)^{-1}=\varepsilon(\sigma),$$

puisque $\mathbb{C}^\times$ est abélien.

En particulier, toutes les transpositions ont même image par ε par le théorème 2.18.

Si $\tau \in \mathfrak{S}(E)$ est une transposition, alors $\tau^2 = \mathrm{Id}_E$, et par conséquent

$$\varepsilon(\tau)^2 = \varepsilon(\tau^2) = \varepsilon(\mathrm{Id}_E) = 1.$$

Ainsi, $\varepsilon(\tau) = \pm 1$. Si l'on avait $\varepsilon(\tau) = 1$ pour toute transposition τ de $\mathfrak{S}(E)$, alors le morphisme ε serait trivial, puisque toute permutation de $\mathfrak{S}(E)$ est produit de transpositions, par le corollaire 2.12. On a donc $\varepsilon(\tau) = -1$ pour toute transposition τ, et donc $\varepsilon(\sigma) = (-1)^s$ si σ est produit de s transpositions. Ainsi, si un tel morphisme ε existe, il est unique.

Montrons l'existence d'un morphisme non trivial $\mathfrak{S}(E) \longrightarrow \mathbb{C}^\times$. La signature (cf. définition IV-1.5) $\varepsilon : \mathfrak{S}_n \longrightarrow \mathbb{C}^\times$ fournit l'existence d'un tel morphisme lorsque $E = [\![1, n]\!]$, d'après la proposition IV-1.6 (3). Choisissons une numérotation $a_1, \ldots, a_n$ des éléments de E et soit

$$u : \mathfrak{S}_n \xrightarrow{\sim} \mathfrak{S}(E)$$

l'isomorphisme correspondant. Alors, $\varepsilon \circ u^{-1} : \mathfrak{S}(E) \longrightarrow \mathbb{C}^\times$ est un morphisme de groupes, non trivial. En effet, si $\sigma \in \mathfrak{S}_n$ vérifie $\varepsilon(\sigma) \neq 1$, alors on a $(\varepsilon \circ u^{-1})\big(u(\sigma)\big) = \varepsilon(\sigma) \neq 1$. Cela achève la démonstration. □

4.2. Remarque. Le théorème donne aussi implicitement le résultat suivant, déjà connu dans le cas de $\mathfrak{S}_n$: la parité du nombre de transpositions nécessaires pour écrire une permutation de $\mathfrak{S}(E)$ ne dépend pas de la décomposition choisie.

4.3. Définition. Si $\sigma \in \mathfrak{S}(E)$, l'élément $\varepsilon_E(\sigma) \in \{\pm 1\}$ est appelé la *signature* de σ.

En général, le morphisme de signature est simplement noté ε.

4.4. Exemple. La signature d'un p-cycle $\sigma = (a_1 \ldots a_p)$ est égale à $(-1)^{p-1}$. En effet, on a

$$(a_1 \ldots a_p) = (a_1\ a_2)(a_2\ a_3) \cdots (a_{p-2}\ a_{p-1})(a_{p-1}\ a_p),$$

et donc σ est produit de $p - 1$ transpositions.

4.5. Proposition. *Soit E un ensemble à n éléments, et soit $\sigma \in \mathfrak{S}(E)$. Alors,*

$$\varepsilon(\sigma) = (-1)^{n - N_\sigma},$$

où N_σ est le nombre de σ-orbites.

Démonstration. Soit $\sigma \in \mathfrak{S}(E)$, et soit $\sigma = \sigma_1 \cdots \sigma_r$ sa décomposition en produit de cycles à supports disjoints. Si l'on note p_i la longueur du cycle σ_i,

alors $p_1 + \cdots + p_r$ est le nombre d'éléments de E qui ne sont pas fixes. Il y a donc $n - (p_1 + \cdots + p_r)$ points fixes, i.e. $n - (p_1 + \cdots + p_r)$ σ-orbites réduites à un élément. Il y a donc $N_\sigma = n - (p_1 + \cdots + p_r) + r$ σ-orbites distinctes, puisque les σ-orbites non réduites à un élément sont les supports des cycles σ_i. D'autre part, on a d'après l'exemple précédent

$$\varepsilon(\sigma) = \varepsilon(\sigma_1) \cdots \varepsilon(\sigma_r) = (-1)^{p_1 - 1} \cdots (-1)^{p_r - 1} = (-1)^{p_1 + \cdots + p_r - r},$$

et donc $\varepsilon(\sigma) = (-1)^{n - N_\sigma}$. Cela achève la démonstration. □

Nous pouvons maintenant définir le groupe alterné en toute généralité.

4.6. Définition. Le *groupe alterné*, noté $\mathfrak{A}(E)$, est l'ensemble des permutations de $\mathfrak{S}(E)$ de signature 1. C'est un sous-groupe distingué de $\mathfrak{S}(E)$ (puisque c'est le noyau d'un morphisme de groupes ; cf. le lemme VI-2.15). En fait, $\mathfrak{A}(E)$ est un sous-groupe d'indice 2 de $\mathfrak{S}(E)$.

En effet, si τ est une transposition fixée, alors pour tout $\sigma \in \mathfrak{S}(E)$, on a soit $\sigma \in \mathfrak{A}(E)$, soit $\tau^{-1}\sigma \in \mathfrak{A}(E)$, et donc $\mathfrak{A}(E)$ et $\tau\mathfrak{A}(E)$ sont les seules classes à gauche de $\mathfrak{S}(E)$ modulo $\mathfrak{A}(E)$.

Si E a n éléments ($n \geqslant 2$) , on a donc

$$|\mathfrak{A}(E)| = \frac{n!}{2}.$$

Si $E = [\![1, n]\!]$, on retrouve le groupe $\mathfrak{A}_n$ introduit dans la définition IV-1.7.

4.7. Remarque. Si $n = 2$, $\mathfrak{A}(E)$ est trivial, et si $n = 3$, $\mathfrak{A}(E)$ est cyclique, engendré par un 3-cycle.

Le résultat suivant, bien qu'il ne nous serve pas pour la suite, est classique.

4.8. Proposition. *Soit E un ensemble ayant au moins deux éléments. Alors, le groupe alterné $\mathfrak{A}(E)$ est l'unique sous-groupe d'indice 2 de $\mathfrak{S}(E)$.*

Démonstration. Soit H un sous-groupe d'indice 2 de $\mathfrak{S}(E)$. On a alors deux classes à gauche H et $\sigma_0 H$, avec $\sigma_0 \notin H$. Toute permutation $\sigma \in \mathfrak{S}(E)$ s'écrit donc de manière unique sous la forme

$$\sigma = \sigma_0^m h, \quad m \in \{0, 1\},\ h \in H.$$

On définit alors une application $f : \mathfrak{S}(E) \longrightarrow \mathbb{C}^\times$ en posant

$$f(\sigma_0^m h) = (-1)^m.$$

Montrons à présent que c'est un morphisme de groupes.

permutation de type $(2,2)$. Il est donc en particulier distingué dans $\mathfrak{A}(E)$. Ainsi, $\mathfrak{A}(E)$ n'est pas simple.

On suppose jusqu'à la fin que $n \geqslant 5$.

Soit H un sous-groupe distingué de $\mathfrak{A}(E)$, non réduit à l'identité. On veut montrer que $H = \mathfrak{A}(E)$. Faisons d'abord quelques remarques préliminaires.

(1) Si H contient un 3-cycle, alors $H = \mathfrak{A}(E)$.

En effet, puisque H contient un 3-cycle, il les contient tous puisque H est distingué dans $\mathfrak{A}(E)$ et que tous les 3-cycles sont conjugués par le lemme précédent. Le groupe $\mathfrak{A}(E)$ étant engendré par les 3-cycles, vu la proposition 4.9, on obtient $H = \mathfrak{A}(E)$.

(2) Si $\sigma \in H$, $\tau \in \mathfrak{A}(E)$, alors $\sigma\tau\sigma^{-1}\tau^{-1} \in H$.

En effet, comme H est un sous-groupe, on a $\sigma^{-1} \in H$, et comme il est distingué dans $\mathfrak{A}(E)$, on a $\tau\sigma^{-1}\tau^{-1} \in H$. On a alors aussi $\sigma\tau\sigma^{-1}\tau^{-1} \in H$.

(3) Si H contient un produit de deux transpositions à supports disjoints, alors H contient un 3-cycle. En particulier, $H = \mathfrak{A}(E)$.

En effet, supposons que l'on ait $\sigma = (a\ b)(c\ d) \in H$, où a, b, c, $d \in E$ sont distincts. Puisque $n \geqslant 5$, on peut choisir un élément $e \in E \setminus \{a, b, c, d\}$. Posons $\tau = (a\ b\ e)$, et remarquons que l'on a $\sigma^{-1} = (c\ d)(a\ b)$. On a alors

$$\tau\sigma^{-1}\tau^{-1} = (\tau(c\ d)\tau^{-1})(\tau(a\ b)\tau^{-1}) = \big(\tau(c)\ \tau(d)\big)\big(\tau(a)\ \tau(b)\big) = (c\ d)(b\ e).$$

Ainsi,
$$\sigma\tau\sigma^{-1}\tau^{-1} = (a\ b)(c\ d)(c\ d)(b\ e) = (a\ b)(b\ e) = (a\ b\ e).$$

Par (2), $(a\ b\ e) \in H$, et par (1), on obtient $H = \mathfrak{A}(E)$.

On tire de tout cela qu'il suffit de montrer que H contient un 3-cycle ou un produit de deux transpositions à supports disjoints, ce que l'on va faire maintenant. Soit $\sigma \in H \setminus \{\mathrm{Id}_E\}$.

Supposons en premier lieu que tous les cycles à supports disjoints composant σ soient des 3-cycles. Si σ est un 3-cycle, il n'y a rien à faire. Sinon, il y en a au moins deux, disons $(a\ b\ c)$ et $(d\ e\ f)$. Soit $\tau = (a\ d)(b\ e)$. Alors, la permutation $\sigma' = \sigma\tau\sigma^{-1}\tau^{-1}$ est dans H, par l'alinéa (2). On a de plus

$$\sigma\tau\sigma^{-1} = \big(\sigma(a)\ \sigma(d)\big)\big(\sigma(b)\ \sigma(e)\big) = (b\ e)(c\ f),$$

et alors

$$\sigma' = (b\ e)(c\ f)(b\ e)(a\ d) = (c\ f)(a\ d) \in H.$$

Supposons maintenant que la décomposition en cycles à supports disjoints de σ ne contienne pas uniquement des cycles d'ordre 3.

Admettons provisoirement l'existence d'une partie F de E à trois éléments telle que $F \cup \sigma(F)$ contienne exactement quatre éléments (en particulier, on a $\sigma(F) \neq F$), et soit τ un cycle de support F. C'est donc un 3-cycle, et

ainsi $\tau \in \mathfrak{A}(E)$. On a de plus

$$\mathrm{Supp}(\sigma\tau\sigma^{-1}) = \sigma(F) \neq F = \mathrm{Supp}(\tau).$$

En particulier, $\sigma\tau\sigma^{-1} \neq \tau$ et donc $\sigma' = \sigma\tau\sigma^{-1}\tau^{-1} \neq \mathrm{Id}_E$. Par (2), on a alors $\sigma' \in H$. De plus, par le lemme 1.5 (3), on a

$$\mathrm{Supp}(\sigma') \subset \mathrm{Supp}(\sigma\tau\sigma^{-1}) \cup \mathrm{Supp}(\tau^{-1}) = \sigma(F) \cup F.$$

Comme $\sigma(F) \cup F$ a quatre éléments et que $\sigma' \neq \mathrm{Id}_E$, σ' est soit un 3-cycle, soit un produit de deux transpositions à supports disjoints, car $\sigma' \in \mathfrak{A}(E)$ (en effet, σ' ne peut être une transposition ou un 4-cycle, car de telles permutations sont de signature -1).

Il reste donc à vérifier l'existence d'une telle partie F. Si au moins un des cycles qui composent σ est de longueur $\geqslant 4$, disons $(a\ b\ c\ d\ldots)$, il suffit de poser $F = \{a, b, c\}$, puisqu'alors

$$F \cup \sigma(F) = \{a, b, c, d\}.$$

Supposons maintenant que tous les cycles qui composent σ soient de longueur 2 ou 3. Par hypothèse, il existe au moins un cycle de longueur 2 dans la décomposition de σ, donc il en existe au moins deux, disons $(a\ b)$ et $(c\ d)$. Dans le cas contraire, la signature de σ serait en effet égale à -1 (la signature d'un 2-cycle est -1 et celle d'un 3-cycle est 1), ce qui est impossible puisque $\sigma \in H \subset \mathfrak{A}(E)$. On prend alors $F = \{a, b, c\}$.

Cela montre le résultat voulu. □

5.4. Remarque. On peut démontrer que $\mathfrak{A}_5$ est l'unique groupe simple d'ordre 60, à isomorphisme près (cf. exercice XI-8).

On va maintenant déduire du théorème 5.3 la liste des sous-groupes distingués de $\mathfrak{S}(E)$. On commence par un lemme.

5.5. Lemme. *Soit E un ensemble à n éléments. Si $n \geqslant 3$, le centre du groupe symétrique $\mathfrak{S}(E)$ est trivial.*

Démonstration. Soit $\sigma \neq \mathrm{Id}_E$, et soit $a \in \mathrm{Supp}(\sigma)$ (le support de σ n'est pas vide puisque $\sigma \neq \mathrm{Id}_E$). On a donc $b = \sigma(a) \neq a$. Puisque $n \geqslant 3$, on peut choisir $c \in E \setminus \{a, b\}$. Soit $\tau = (b\ c)$. On a alors

$$\sigma\big(\tau(a)\big) = \sigma(a) = b \quad \text{et} \quad \tau\big(\sigma(a)\big) = \tau(b) = c.$$

Ainsi, σ et τ ne commutent pas, et donc $\sigma \notin Z\big(\mathfrak{S}(E)\big)$. Par conséquent, on a $Z\big(\mathfrak{S}(E)\big) = \{\mathrm{Id}_E\}$. □

5.6. Remarque. Le résultat est faux pour $n = 2$, puisque $\mathfrak{S}_2$ est commutatif.

5.7. Corollaire. *Soit E un ensemble à n éléments ($n \geqslant 2$). Si $n \neq 4$, les sous-groupes distingués de $\mathfrak{S}(E)$ sont*

$$\{\mathrm{Id}_E\}, \quad \mathfrak{A}(E) \quad \textit{et} \quad \mathfrak{S}(E).$$

Démonstration. Si $n = 2$, le résultat est clair puisque $\mathfrak{A}(E)$ est réduit à l'identité. On peut donc supposer que $n \geqslant 3$ et $n \neq 4$. Soit H un sous-groupe distingué de $\mathfrak{S}(E)$. Alors, $H \cap \mathfrak{A}(E)$ est un sous-groupe distingué de $\mathfrak{A}(E)$, et donc $H \cap \mathfrak{A}(E) = \{\mathrm{Id}_E\}$ ou $\mathfrak{A}(E)$ par le théorème 5.3.

Si $H \cap \mathfrak{A}(E) = \mathfrak{A}(E)$, alors $\mathfrak{A}(E) \subset H$. Si $H \subset \mathfrak{A}(E)$, alors $H = \mathfrak{A}(E)$. Si H contient une permutation σ_0 de signature -1, il contient alors deux classes à gauche modulo $\mathfrak{A}(E)$ distinctes, à savoir $\mathfrak{A}(E)$ et $\sigma_0\mathfrak{A}(E)$, c'est-à-dire qu'il contient $\mathfrak{S}(E)$. Dans ce cas $H = \mathfrak{S}(E)$.

Si $H \cap \mathfrak{A}(E) = \{\mathrm{Id}_E\}$, alors la signature induit par restriction un morphisme injectif $H \hookrightarrow \{\pm 1\}$ (car la signature d'une permutation vaut ± 1), de sorte que $|H| \leqslant 2$.

Supposons que $|H| = 2$, et soit $\tau \in H$ l'unique élément non trivial de H. Soit $\sigma \in \mathfrak{S}(E)$. Comme $\sigma\tau\sigma^{-1} \in H$ et $\sigma\tau\sigma^{-1} \neq \mathrm{Id}_E$ (puisque $\tau \neq \mathrm{Id}_E$), on a $\sigma\tau\sigma^{-1} = \tau$. On en déduit que $\tau \in Z\big(\mathfrak{S}(E)\big)$. Comme $n \geqslant 3$, le centre de $\mathfrak{S}(E)$ est trivial d'après le lemme précédent, et donc $\tau = \mathrm{Id}_E$, d'où une contradiction. Ainsi, $|H| = 1$, et donc $H = \{\mathrm{Id}_E\}$ dans ce cas. Cela achève la démonstration. □

5.8. Remarque. Le résultat est faux si $n = 4$. En effet, si $E = \{a, b, c, d\}$, l'ensemble

$$H = \{\mathrm{Id}_E,\ (a\ b)(c\ d),\ (a\ c)(b\ d),\ (a\ d)(b\ c)\}$$

est un sous-groupe distingué de $\mathfrak{S}(E)$, distinct de $\{\mathrm{Id}_E\}$, $\mathfrak{A}(E)$ et $\mathfrak{S}(E)$.

On peut d'ailleurs montrer que les sous-groupes distingués de $\mathfrak{S}(E)$ sont dans ce cas

$$\{\mathrm{Id}_E\}, \quad H, \quad \mathfrak{A}(E) \quad \text{et} \quad \mathfrak{S}(E).$$

6. Exercices

E1. Dans $\mathfrak{S}_{11}$, on considère

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 3 & 2 & 6 & 4 & 1 & 5 & 7 & 11 & 9 & 10 & 8 \end{pmatrix}$$

et

$$\tau = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 1 & 9 & 3 & 4 & 5 & 6 & 2 & 8 & 10 & 7 & 11 \end{pmatrix}.$$

Montrer sans calculs que σ et τ commutent, puis le vérifier par calcul direct.

E2. Décomposer en produit de cycles à supports deux à deux disjoints les permutations de $\mathfrak{S}_{10}$ suivantes :

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 5 & 8 & 2 & 7 & 9 & 6 & 10 & 3 & 1 & 4 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 3 & 5 & 4 & 1 & 7 & 10 & 2 & 6 & 9 & 8 \end{pmatrix},$$

$$(1\ 2\ 3\ 4\ 5)(1\ 5\ 9)(2\ 4\ 9)\ ,\ (1\ 8\ 6\ 7)(8\ 6\ 7\ 5)(6\ 7\ 5\ 1),$$

$$(1\ 2)(2\ 3)(3\ 4)(4\ 5)(5\ 6)(6\ 7)(7\ 1)\ ,\ (1\ 8)(1\ 7)(1\ 6)(1\ 5)(1\ 4)(1\ 3)(1\ 2).$$

E3. Trouver un 4-cycle σ tel que $(1\ 2)$ et σ n'engendrent pas $\mathfrak{S}_4$ (penser aux isométries du carré).

E4. En l'an 3000, il existe une machine qui permet à deux personnes d'échanger leurs corps, une et une seule fois (on ne peut pas revenir en arrière).

Montrer que si n personnes ont utilisé la machine (éventuellement plusieurs fois) pour échanger leurs corps, alors ces n personnes peuvent réintégrer leur propre corps en utilisant deux personnes supplémentaires (qui n'ont pas utilisé la machine) [2].

Indication. Commencer par le cas d'un cycle.

E5. Soit $\sigma \in \mathfrak{S}_5$ la permutation $\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 5 & 1 & 3 \end{pmatrix}$.

Calculer la puissance σ^{2017}.

E6. Calculer l'ordre de chacune des permutations de l'exercice 2.

E7. Un joueur professionnel mélange un paquet de cartes plusieurs fois par la méthode suivante : il sépare le paquet en deux moitiés égales, puis

2. C'est le théorème de Futurama, dû à Ken Keeler, scénariste de l'excellente série du même nom.

intercale une carte sur deux de chaque paquet, en gardant la première carte de la première moitié sur le dessus.

L'expérimentation avec huit cartes montre qu'au bout de trois mélanges successifs, les cartes sont revenues dans le même ordre ! S'agit-il d'un cas particulier ? Le même phénomène se produit-il avec un nombre arbitraire pair de cartes ? Il pourra sans doute être commode de noter les cartes de 0 à $2n-1$.

Que se passe-t-il si l'on change la façon de mélanger en gardant la première carte de la seconde moitié sur le dessus ?

E8. Les permutations suivantes sont-elles conjuguées ? Si oui, donner une permutation qui les conjugue.

(1) $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 5 & 8 & 2 & 7 & 9 & 6 & 10 & 3 & 1 & 4 \end{pmatrix}$ et $\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 \\ 3 & 5 & 4 & 1 & 7 & 10 & 2 & 6 & 9 & 8 \end{pmatrix}$

(2) $(1\,2\,3\,4\,5)(1\,5\,9)(2\,4\,9)$ et $(1\,8\,6\,7)(8\,6\,7\,5)(6\,7\,5\,1)$

(3) $(1\,2)(2\,3)(3\,4)(4\,5)(5\,6)(6\,7)(7\,1)$ et
$(1\,8)(1\,7)(1\,6)(1\,5)(1\,4)(1\,3)(1\,2)$.

E9. Donner les classes de conjugaison de $\mathfrak{S}_5$, l'ordre maximal d'un élément de $\mathfrak{S}_{10}$, les points fixes et les stabilisateurs pour l'action par conjugaison de $\mathfrak{S}_3$ (resp. $\mathfrak{S}_4$) sur lui-même, et le nombre de classes de conjugaison des éléments d'ordre 6 dans $\mathfrak{S}_9$.

E10. Calculer la signature de chaque permutation de l'exercice 2.

E11. Déterminer tous les morphismes de $\mathfrak{S}_n$ dans $\mathbb{Z}/2\mathbb{Z}$, puis $\mathbb{Z}/4\mathbb{Z}$, puis dans tout groupe abélien A d'ordre impair.

E12. On propose dans cet exercice une définition alternative de la signature d'une permutation de $\mathfrak{S}_n$. On dénote par X l'ensemble des paires d'éléments $\{i,j\}$ $(i \neq j)$ de l'ensemble $[\![1,n]\!]$.

a. Soit $\sigma \in \mathfrak{S}_n$. Montrer que le rationnel $\dfrac{\sigma(i)-\sigma(j)}{i-j}$ ne dépend que de la paire $\{i,j\}$.

b. Pour $\sigma \in \mathfrak{S}_n$, on pose

$$\varphi(\sigma) = \prod_{\{i,j\} \in X} \frac{\sigma(i) - \sigma(j)}{i - j} \in \mathbb{C}^\times.$$

Montrer que $\varphi : \mathfrak{S}_n \longrightarrow \mathbb{C}^\times$ est un morphisme de groupes.

c. Calculer $\varphi\big((1\ 2)\big)$, et en déduire que φ est le morphisme de signature.

E13. Si $n \geqslant 3$, montrer que $\mathfrak{A}_n$ est engendré par chacune des familles suivantes :

(1) les cycles de la forme $(1\ i\ j)$, $i < j$;

(2) les 3-cycles $(1\ 2\ i)$, $i \in [\![3, n]\!]$;

(3) les 3-cycles $(i\ i+1\ i+2)$, $i \in [\![1, n-2]\!]$.

E14. Trouver deux 3-cycles non conjugués dans $\mathfrak{A}_n$ pour $n = 3, 4$.

E15. Soit $n \geqslant 3$. Montrer que $\mathfrak{A}_n$ est engendré par les éléments σ^2, $\sigma \in \mathfrak{S}_n$.

Indication. Si σ est un 3-cycle, que vaut σ^4 ?

E16.

a. Trouver toutes les formes possibles de décomposition en cycles à supports deux à deux disjoints des éléments de $\mathfrak{A}_4$.

b. En déduire qu'un élément de $\mathfrak{A}_4$ est d'ordre $1, 2$ ou 3, et déterminer le nombre d'éléments de $\mathfrak{A}_4$ d'un ordre donné.

c. On veut montrer que $\mathfrak{A}_4$ n'a pas de sous-groupe d'ordre 6. On suppose que H est un tel sous-groupe.

Justifier que H est distingué dans $\mathfrak{A}_4$, et en déduire que $\sigma^2 \in H$, pour tout $\sigma \in \mathfrak{A}_4$. Conclure en utilisant l'exercice précédent.

d. Montrer que les seuls sous-groupes distingués de $\mathfrak{A}_4$ sont $\{\mathrm{Id}\}, \mathfrak{A}_4$ et le groupe engendré par les éléments de type $(2, 2)$.

E17. Montrer que toute permutation d'ordre 10 dans $\mathfrak{S}_8$ est de signature égale à -1.

E18. Existe-t-il un morphisme injectif de $\mathfrak{S}_7$ dans $\mathfrak{A}_8$?

E19. Soit p un nombre premier. Montrer que les éléments d'ordre p du groupe $\mathfrak{S}_p$ sont exactement les p-cycles. Le résultat subsiste-t-il si p n'est pas premier ?

E20. Soit $\sigma \in \mathfrak{S}_n$ un k-cycle.

On veut démontrer le résultat suivant : pour tout $d \geqslant 1$, σ^d se décompose en $\operatorname{pgcd}(k,d)$ cycles de longueur $\dfrac{k}{\operatorname{pgcd}(k,d)}$ à supports deux à deux disjoints.

a. Montrer qu'il suffit d'établir le résultat pour $k = n$.

b. Calculer le stabilisateur de $i \in [\![1, n]\!]$ sous l'action de $\langle \sigma^d \rangle$.

c. Conclure.

d. En déduire les deux cas particuliers suivants : si $\sigma \in \mathfrak{S}_n$ est un k-cycle, et si $d \geqslant 1$ est un entier, alors :

▷ si $d \mid k$, σ^d se décompose en d cycles de longueur $\dfrac{k}{d}$ à supports deux à deux disjoints

▷ si d et k sont premiers entre eux, σ^d est un k-cycle.

e. Le résultat subsiste-t-il si l'on remplace $\mathfrak{S}_n$ par $\mathfrak{S}(E)$, où E est un ensemble à n éléments ?

f. Soit $\sigma \in \mathfrak{S}_{20}$ de type $(8, 6, 3, 1, 1, 1)$. Calculer le type de σ^4.

E21. Soit p un nombre premier. On veut démontrer qu'une transposition et un p-cycle engendrent $\mathfrak{S}_p$. Soit σ un p-cycle et soit τ une transposition.

a. Soit $\rho \in \mathfrak{S}_p$. Montrer que σ et τ engendrent $\mathfrak{S}_p$ si, et seulement si, les permutations $\rho\sigma\rho^{-1}$ et $\rho\tau\rho^{-1}$ engendrent $\mathfrak{S}_p$. En déduire que l'on peut supposer $\tau = (1\ 2)$.

Indication. Remarquer que $\operatorname{Int}(\rho)$ est un automorphisme de G.

b. Montrer qu'il existe $d \in [\![1, p-1]\!]$ tel que $\sigma^d(1) = 2$. Justifier que σ^d est encore un p-cycle.

Indication. On pourra utiliser l'exercice 19 ou l'exercice 20.

c. Montrer que σ et τ engendrent $\mathfrak{S}_p$ si, et seulement si, σ^d et τ engendrent $\mathfrak{S}_p$. En déduire que l'on peut supposer que $\sigma(1) = 2$.

On suppose donc que $\sigma = (1\ 2\ a_3 \dots a_p)$ et $\tau = (1\ 2)$.

d. Montrer qu'il existe $\rho \in \mathfrak{S}_p$ tel que

$$\rho\sigma\rho^{-1} = (1\ 2 \dots p) \quad \text{et} \quad \rho\tau\rho^{-1} = (1\ 2).$$

Conclure.

E22. Quelle est la probabilité qu'une permutation de $2n$ éléments possède un cycle de longueur $> n$ dans sa décomposition en cycles à supports deux à deux disjoints ?

Application. Un directeur de prison veut faire de la place et libérer cent cellules. Il convoque alors cent prisonniers et les invite à jouer au jeu suivant. Il a aligné cent urnes contenant chacune le nom d'un des cent prisonniers (chaque nom figure une et une seule fois). Chaque prisonnier pourra à tour de rôle regarder le contenu de cinquante urnes. S'il ne trouve pas son nom, le prisonnier sera exécuté ainsi que tous ses camarades, sinon le jeu continue. Si à la fin du jeu, tous les prisonniers retrouvent leur nom, ils seront tous libérés. Les prisonniers ont quelques instants pour mettre une stratégie au point, sachant qu'une fois le jeu commencé, ils ne pourront plus communiquer entre eux, ni laisser des marques sur les urnes ou déplacer les urnes ou les messages. En choisissant les urnes au hasard, leur chance de s'en sortir vivants est quasi nulle, mais il existe une stratégie leur donnant plus de 30% de chances de réussir. Quelle est cette stratégie ?

On pourra supposer que les prisonniers se numérotent de 1 à 100 et se souviennent tous du nom de chacun d'entre eux. Les urnes sont aussi numérotées de 1 à 100.

E23. On veut établir le résultat suivant : si $\sigma \in \mathfrak{S}_n$, et si N_σ est le nombre de σ-orbites, alors le nombre minimal de transpositions nécessaires pour écrire σ comme produit de transpositions est $n - N_\sigma$.

a. Montrer qu'un cycle de longueur k ($k \geqslant 2$) peut s'écrire comme produit de $k-1$ transpositions. En déduire qu'il existe une décomposition de σ en produit de $n - N_\sigma$ transpositions.

Soit $(e_1, \ldots, e_n)$ la base canonique de $\mathbb{R}^n$. Pour $\sigma \in \mathfrak{S}_n$, on note f_σ l'unique endomorphisme de $\mathbb{R}^n$ tel que

$$f_\sigma(e_i) = e_{\sigma(i)}, \quad \text{pour tout } i \in [\![1, n]\!].$$

b. Montrer que $f_\sigma \in \mathrm{O}(\mathbb{R}^n)$, et que l'application

$$\begin{aligned} \mathfrak{S}_n &\longrightarrow \mathrm{O}(\mathbb{R}^n) \\ \sigma &\longmapsto f_\sigma \end{aligned}$$

est un morphisme de groupes injectif.

c. Montrer que $\dim_{\mathbb{R}} \mathrm{Ker}(f_\sigma - \mathrm{Id}_{\mathbb{R}^n}) = n - N_\sigma$.

d. On suppose que σ est une transposition. Montrer que f_σ est une réflexion orthogonale.

e. Conclure en utilisant les questions précédentes et le théorème V-3.7.

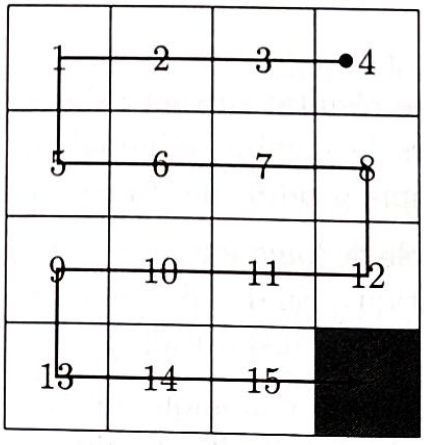
E24. Soit $\sigma \in \mathfrak{S}_n$ une permutation de support I. Montrer que $\tau \in \mathfrak{S}_n$ commute avec σ si, et seulement si, $\tau(I) = I$ et que la permutation τ_I induite par τ sur I est une puissance de celle induite par σ sur I.

E25. On considère le jeu de taquin suivant :

1	2	3	4
5	6	7	8
9	10	11	12
13	15	14	

Peut-on arriver à remettre les pièces dans l'ordre sans démonter le jeu ?

Indication. On pourra associer à une configuration du taquin une permutation de $\mathfrak{S}_{15}$ en suivant le chemin suivant (sans tenir compte de la case noire)



et étudier l'effet d'un déplacement sur la signature de cette permutation [3].

E26. Soit G un groupe fini. On rappelle que l'on a un morphisme de groupes injectif $\varphi : G \longrightarrow \mathfrak{S}(G)$, qui à $g \in G$ associe la permutation

$$\begin{aligned} \varphi(g) : G &\longrightarrow G \\ x &\longmapsto gx. \end{aligned}$$

On se propose de calculer la signature de $\varphi(g)$.

3. Le lecteur pourra trouver un approfondissement sur le jeu de taquin dans [23], dont est tiré le chemin proposé dans l'indication.

a. Soit E l'ensemble des classes **à droite** de G modulo $\langle g \rangle$, et soit $r = |E|$. Justifier que $r = \dfrac{|G|}{o(g)}$.

b. Montrer que toute classe à droite est stable par $\varphi(g)$, et en déduire que $\varphi(g)$ se décompose en produit de r cycles de longueur $o(g)$, de supports deux à deux disjoints.

c. En déduire que $\varepsilon\big(\varphi(g)\big) = (-1)^{(o(g)-1)\frac{|G|}{o(g)}}$.

d. On écrit $|G| = 2^m s$, avec s impair et $m \geqslant 0$. Montrer que $\varphi(G) \subset \mathfrak{A}(G)$ si, et seulement si, $|G|$ est impair ou G est d'ordre pair et ne possède pas d'élément d'ordre 2^m.

e. On reprend les notations de la question **d.** On suppose que G possède un élément d'ordre 2^m. Construire un morphisme non trivial

$$\rho : G \longrightarrow \{\pm 1\}.$$

En déduire que G n'est pas simple si $|G| \geqslant 3$.

E27. On se place dans l'espace euclidien $\mathbb{R}^3$. Soit T l'ensemble des sommets d'un tétraèdre régulier centré en l'origine. Soit G l'ensemble des isométries de $\mathbb{R}^3$ qui laissent T globalement invariant.

a. Montrer que G agit transitivement sur T et étudier le stabilisateur d'un sommet.

b. En déduire que G est fini, et donner son cardinal.

c. Montrer que G est isomorphe à $\mathfrak{S}_4$.

d. On note G^+ l'ensemble des isométries directes laissant T globalement invariant.

(i) Si $s \in G$ est une isométrie indirecte, montrer que tout élément de G est de la forme f ou sf, avec $f \in G^+$. En déduire que $[G : G^+] = 2$.

(ii) Montrer que $G^+ \simeq \mathfrak{A}_4$.

e. Décrire géométriquement tous les éléments de G^+.

f. Écrire l'inventaire des coloriages d'un tétraèdre en au plus q couleurs, et donner le nombre de coloriages possibles (cf. paragraphe VII-3).

E28. On considère un cube C centré en l'origine $O = (0, 0, 0)$. On note G (resp. G^+) le groupe des isométries (resp. le groupes des isométries directes) de $\mathbb{R}^3$ conservant C globalement.

a. Vérifier que $-\mathrm{Id}_{\mathbb{R}^3} \in G$, et que l'application

$$\begin{aligned} G^+ \times \{\pm 1\} &\longrightarrow G \\ (f, \varepsilon) &\longmapsto \varepsilon f \end{aligned}$$

est un isomorphisme de groupes.

b. (i) Vérifier que le groupe G agit sur l'ensemble X des grandes diagonales du cube.

(ii) Vérifier qu'il existe un élément de G qui échange deux grandes diagonales données, et qui fixe globalement les deux autres.

(iii) En déduire l'existence d'un morphisme surjectif $\varphi : G \longrightarrow \mathfrak{S}_4$.

(iv) Justifier que le morphisme $\varphi_{|_{G^+}} : G^+ \longrightarrow \mathfrak{S}_4$ est encore surjectif.

c. Soit $f \in \mathrm{Ker}(\varphi)$. On suppose que $f \neq \mathrm{Id}_{\mathbb{R}^3}$.

(i) Vérifier que f conserve globalement chaque grande diagonale, mais que f échange deux sommets d'au moins une grande diagonale.

(ii) Montrer alors que f échange les sommets de toute grande diagonale, et en déduire que $f = -\mathrm{Id}_{\mathbb{R}^3}$.

d. En déduire que $\mathrm{Ker}(\varphi) = \{\pm \mathrm{Id}_{\mathbb{R}^3}\}$.

e. Déduire des questions précédentes que l'on a des isomorphismes

$$G^+ \simeq \mathfrak{S}_4 \quad \text{et} \quad G \simeq \mathfrak{S}_4 \times \mathbb{Z}/2\mathbb{Z}.$$

f. Décrire géométriquement les éléments de G^+.

g. Écrire l'inventaire des coloriages d'un cube en au plus q couleurs, et donner le nombre de coloriages possibles (cf. paragraphe VII-3).

E29. Soit $\rho \in \mathfrak{S}_n$. On pose $C(\rho) = \{\sigma \in \mathfrak{S}_n \mid \sigma\rho\sigma^{-1} = \rho\}$.

a. Montrer que $C(\rho)$ est un sous-groupe de $\mathfrak{S}_n$.

b. Montrer que $C\big(f(\rho)\big) \simeq C(\rho)$, pour tout automorphisme f de $\mathfrak{S}_n$.

On se propose de calculer $|C(\rho)|$. On suppose que ρ possède k_1 points fixes, et que ρ est le produit de k_2 2-cycles, ..., k_n n-cycles, tous de supports deux à deux disjoints.

c. Soit $X(\rho)$ la classe de conjugaison de ρ.

(i) Montrer que $|X(\rho)| = \dfrac{n!}{|C(\rho)|}$.

(ii) Montrer que le nombre de permutations conjuguées à ρ est

$$\frac{n!}{\prod_{i=1}^{n} k_i!\, i^{k_i}}.$$

(iii) En déduire que $|C(\rho)| = \prod_{i=1}^{n} k_i!\, i^{k_i}$.

E30.

a. Soient E, E' deux ensembles à n éléments ($n \geqslant 2$). Montrer que tout isomorphisme $\varphi : \mathfrak{S}(E) \xrightarrow{\sim} \mathfrak{S}(E')$ induit par restriction un isomorphisme $\mathfrak{A}(E) \simeq \mathfrak{A}(E')$.

Indication. Vérifier qu'un isomorphisme préserve l'indice des sous-groupes.

Soit E un ensemble à n éléments, et soit $a \in E$. Soit $F = E \setminus \{a\}$.

b. Soit $H = \{\sigma \in \mathfrak{S}(E) \mid \sigma(a) = a\}$. Montrer que l'on a des isomorphismes

$$H \simeq \mathfrak{S}(F) \simeq \mathfrak{S}_{n-1}.$$

c. On suppose que $n \geqslant 3$. Soit $H' = \{\sigma \in \mathfrak{A}(E) \mid \sigma(a) = a\}$. Montrer que l'on a des isomorphismes

$$H' \simeq \mathfrak{A}(F) \simeq \mathfrak{A}_{n-1}.$$

d. On suppose que $n \neq 4$. Soit G un sous-groupe de $\mathfrak{S}(E)$ d'indice n.

(i) On fait agir $\mathfrak{S}(E)$ sur $X = \mathfrak{S}(E)/G$ par translation. Montrer que le morphisme $\varphi : \mathfrak{S}(E) \longrightarrow \mathfrak{S}(X)$ est non trivial, et que son image contient au moins n éléments distincts (utiliser la transitivité de cette action).

(ii) Montrer que φ est injectif. On étudiera l'indice de $[\mathfrak{S}(E) : \mathrm{Ker}(\varphi)]$.

(iii) Montrer que pour tout $g \in G$, $\varphi(g)$ fixe $\overline{\mathrm{Id}}_E \in X$.

(iv) En déduire que $G \simeq \mathfrak{S}_{n-1}$.

e. On suppose que $n \neq 4$. Soit G un sous-groupe d'indice n de $\mathfrak{A}(E)$.

(i) On fait agir $\mathfrak{A}(E)$ sur $X' = \mathfrak{A}(E)/G$ par translation. Montrer que le morphisme $\varphi : \mathfrak{A}(E) \longrightarrow \mathfrak{S}(X')$ est non trivial.

(ii) Montrer que φ est injectif.

(iii) En composant φ par le morphisme de signature, montrer que φ induit un morphisme injectif

$$\varphi : \mathfrak{A}(E) \longrightarrow \mathfrak{A}(X').$$

(iv) En déduire comme précédemment que $G \simeq \mathfrak{A}_{n-1}$.

f. Examiner le cas $n = 4$.

E31. On se propose d'étudier ici les automorphismes de $\mathfrak{S}_n$, $n \neq 6$.

Pour tout $i \in [\![2, n]\!]$, on pose $\tau_i = (1\ i)$. On rappelle que $\tau_2, \ldots, \tau_n$ engendrent $\mathfrak{S}_n$ pour $n \geqslant 2$. Soit $f \in \mathrm{Aut}(\mathfrak{S}_n)$.

Le but des questions suivantes est de montrer que f est un automorphisme intérieur de $\mathfrak{S}_n$ si, et seulement si, l'image de toute transposition par f est une transposition.

a. On suppose que f est un automorphisme intérieur de $\mathfrak{S}_n$. Montrer que pour toute transposition $\tau \in \mathfrak{S}_n$, $f(\tau)$ est encore une transposition.

b. On suppose maintenant que f envoie toute transposition sur une transposition. En particulier, on a $f(\tau_i) = (a_i\ b_i)$.

(i) Montrer que si $i \neq j$, alors $\{a_i, b_i\} \cap \{a_j, b_j\} \neq \emptyset$

Indication. Dans le cas contraire, montrer que τ_i et τ_j commutent.

(ii) On suppose, quitte à échanger a_3 et b_3, que $a_2 = a_3$.
Montrer alors par l'absurde que a_2 appartient au support de $f(\tau_i)$, pour tout entier $i \in [\![2, n]\!]$.

Indication. Si a_2 n'est pas dans le support de $f(\tau_i)$, justifier que l'on a $i \geqslant 4$ et $f(\tau_i) = (b_2\ b_3)$, puis considérer $f(\tau_2)f(\tau_3)f(\tau_i)$.

Quitte à échanger certains éléments, on peut donc toujours supposer que l'on a $f(\tau_i) = (a_2\ b_i)$, $a_2 \neq b_i$, pour tout $i \in [\![2, n]\!]$.

(iii) Justifier que $\{a_2, b_2, \ldots, b_n\} = \{1, \ldots, n\}$.

(iv) On pose $b_1 = a_2$. Soit $\rho \in \mathfrak{S}_n$ définie par $\rho(k) = b_k$, $k \in [\![1, n]\!]$. Calculer $\rho\tau_i\rho^{-1}$, et en déduire que f est un automorphisme intérieur.

c. Soit $\tau \in \mathfrak{S}_n$ une transposition et soit $f \in \mathrm{Aut}(\mathfrak{S}_n)$ un automorphisme arbitraire.

(i) Justifier que $f(\tau)$ est d'ordre 2, et s'écrit comme produit de k transpositions à supports deux à deux disjoints, pour un certain $k \geqslant 1$.

(ii) Déduire de l'exercice 29 que $2(n-2)! = 2^k k!(n-2k)!$.

(iii) Montrer que si $n = 2, 3$, on a nécessairement $k = 1$.

(iv) Montrer que si $n = 4, 5$, l'égalité du (2) n'est vraie que pour $k = 1$.

(v) On suppose que $n \geqslant 7$. Montrer que $k! \leqslant (n-k)\cdots(n-2k+1)$.
Montrer également que si $k \geqslant 3$, on a

$$2^{k-1} < (n-2)\cdots(n-k+1).$$

(vi) En déduire que si $n \geqslant 7$, on a $k = 1$.

(vii) Conclure.

E32. On se propose ici de montrer que les automorphismes de $\mathfrak{S}_6$ ne sont pas tous intérieurs.

a. Vérifier qu'il y a quinze transpositions dans $\mathfrak{S}_6$.

On appelle *triple-transposition* un produit de trois transpositions à supports deux à deux disjoints. On appelle *pentade* de $\mathfrak{S}_6$ un ensemble P de cinq triple-transpositions tel que chaque transposition de $\mathfrak{S}_6$ apparaissent dans une et une seule triple-transposition de P (autrement dit, on sépare les transpositions en cinq paquets de trois transpositions, et l'on fait le produit des transpositions de chaque paquet).

b. Soit P une pentade fixée.

(i) Remarquer que les transpositions (1 2), (1 3) et (2 3) sont contenues dans des éléments de P deux à deux distincts.

(ii) En déduire que P contient un élément de la forme

$$\big(1\ \rho(4)\big)\big(2\ \rho(5)\big)\big(3\ \rho(6)\big),$$

où ρ est une permutation de $\{4,5,6\}$. En déduire qu'il y a au plus six pentades distinctes.

(iii) Montrer que le dernier élément de P est alors de la forme

$$\big(1\ \rho\sigma^\ell(4)\big)\big(2\ \rho\sigma^\ell(5)\big)\big(3\ \rho\sigma^\ell(6)\big),$$

où $\sigma = (4\ 5\ 6)$ et $\ell \in \{1,2\}$.

(iv) On pose $\tau_1 = (4\ 5)$, $\tau_2 = (4\ 6)$, $\tau_3 = (5\ 6)$. Justifier qu'il existe une unique pentade contenant

$$\big(1\ \tau_k(4)\big)\big(2\ \tau_k(5)\big)\big(3\ \tau_k(6)\big) \quad \text{et} \quad \big(1\ \tau_k\sigma^\ell(4)\big)\big(2\ \tau_k\sigma^\ell(5)\big)\big(3\ \tau_k\sigma^\ell(6)\big),$$

et expliciter ses éléments.

(v) En déduire qu'il y a exactement six pentades.

c. Montrer que l'action par conjugaison de $\mathfrak{S}_6$ sur lui-même induit une action de $\mathfrak{S}_6$ sur l'ensemble E des pentades.

Pour $k \in [\![1,3]\!]$ et $\ell \in \{1,2\}$, on note $P_{2(k-1)+\ell}$ la pentade de la question **b**, (iv). On obtient ainsi une numérotation des éléments de E.

d. En déduire l'existence d'un morphisme de groupes $\varphi : \mathfrak{S}_6 \longrightarrow \mathfrak{S}_6$ tel que $\varphi\big((1\ 2\ 3)\big) \neq \mathrm{Id}$.

e. Montrer que φ est injectif, puis un automorphisme de $\mathfrak{S}_6$.

f. Calculer $\varphi\big((1\ 2)\big)$ et en déduire que φ n'est pas un automorphisme intérieur de $\mathfrak{S}_6$.

E33. On se propose ici d'étudier les classes de conjugaison de $\mathfrak{A}(E)$.

Si $\sigma \in \mathfrak{A}(E)$, on note $X_s(\sigma)$ la classe de conjugaison de σ dans $\mathfrak{S}(E)$, et $X_a(\sigma)$ la classe de conjugaison de σ dans $\mathfrak{A}(E)$. On note aussi

$$C(\sigma) = \{\rho \in \mathfrak{S}(E) \mid \rho\sigma = \sigma\rho\}.$$

On remarque que $X_a(\sigma) \subset X_s(\sigma)$.

a. Soit $\sigma \in \mathfrak{A}(E)$.

(i) On suppose que $\tau_1\sigma\tau_1^{-1} = \tau_2\sigma\tau_2^{-1}$. Montrer qu'il existe $\rho \in C(\sigma)$ tel que $\tau_2 = \tau_1\rho$.

(ii) On suppose que $C(\sigma)$ n'est pas contenu dans $\mathfrak{A}(E)$. Montrer que l'on a $X_s(\sigma) = X_a(\sigma)$.

(iii) On suppose que $C(\sigma) \subset \mathfrak{A}(E)$. Montrer que $X_a(\sigma) \subsetneq X_s(\sigma)$. Soit $\tau \in \mathfrak{S}(E)$ tel que $\tau\sigma\tau^{-1}$ ne soit pas conjugué à σ dans $\mathfrak{A}(E)$. Montrer que $X_s(\sigma)$ est la réunion disjointe de $X_a(\sigma)$ et de $X_a(\tau\sigma\tau^{-1})$.

Soit $\sigma = \sigma_1 \cdots \sigma_r$ la décomposition de σ en cycles à supports deux à deux disjoints.

b.

(i) On suppose qu'il existe i tel que σ_i soit un cycle de longueur paire. En remarquant que $\sigma_i \in C(\sigma)$, montrer que $X_s(\sigma) = X_a(\sigma)$.

(ii) On suppose que σ admet au moins deux points fixes a et b. Montrer que $(a\ b) \in C(\sigma)$ et en déduire que $X_s(\sigma) = X_a(\sigma)$.

(iii) On suppose qu'il existe $i \neq j$ tels que

$$\sigma_i = (a_1 \dots a_p) \quad \text{et} \quad \sigma_j = (b_1 \dots b_p)$$

soient deux cycles de même longueur impaire.

Vérifier que $(a_1\ b_1) \cdots (a_p\ b_p)$ est un élément de $C(\sigma)$, et montrer que $X_s(\sigma) = X_a(\sigma)$.

(iv) On suppose que les σ_i sont tous de longueurs différentes et impaires, et que σ possède au plus un point fixe. Montrer l'égalité

$$C(\sigma) = \langle \sigma_1, \dots, \sigma_r \rangle,$$

puis que $C(\sigma) \subset \mathfrak{A}(E)$. En déduire que $X_s(\sigma)$ se scinde en deux classes de conjugaison dans $\mathfrak{A}(E)$.

Indication. On pourra d'abord montrer par récurrence sur r que l'on a $|\langle \sigma_1, \dots, \sigma_r \rangle| = |\langle \sigma_1 \rangle| \cdots |\langle \sigma_r \rangle|$, et utiliser l'exercice 29.

c. Résumer le résultat obtenu.

d. Soient $\sigma, \sigma' \in \mathfrak{A}(E)$. On suppose que σ et σ' sont conjuguées dans le groupe $\mathfrak{S}(E)$. Soit $\tau \in \mathfrak{S}(E)$ une permutation telle que $\sigma' = \tau\sigma\tau^{-1}$.

(i) Montrer que pour tout $\rho \in \mathfrak{S}(E)$, on a $\sigma' = \rho\sigma\rho^{-1}$ si, et seulement si, $\rho = \tau\nu$, avec $\nu \in C(\sigma)$.

(ii) On suppose que σ possède dans sa composition un cycle de longueur paire, ou au moins deux cycles de même longueur impaire, ou encore au moins deux points fixes. Donner un algorithme pour trouver un élément $\rho \in \mathfrak{A}(E)$ qui conjugue σ en σ'.

(iii) On suppose que σ est composée de cycles de longueurs impaires deux à deux distinctes et possède au plus un point fixe. Montrer que les permutations σ et σ' sont conjuguées dans $\mathfrak{A}(E)$ si, et seulement si, on a $\tau \in \mathfrak{A}(E)$.

e. Détailler un algorithme qui décide si deux permutations de $\mathfrak{A}(E)$ sont conjuguées dans $\mathfrak{A}(E)$, et qui explicite un élément de $\mathfrak{A}(E)$ qui les conjugue le cas échéant.

f. Application numérique

(1) $\sigma = (1\ 2\ 3\ 4\ 5)$, $\sigma' = (1\ 2\ 3\ 5\ 4)$ dans $\mathfrak{A}_5$.

(2) $\sigma = (1\ 2)(3\ 4)(5\ 7\ 6)$, $\sigma' = (1\ 7)(3\ 5)(2\ 4\ 6)$ dans $\mathfrak{A}_7$.

(3) $\sigma = (1\ 2\ 3)(4\ 5\ 6)$, $\sigma' = (1\ 4\ 2)(3\ 6\ 5)$ dans $\mathfrak{A}_6$.

g. Énumérer les classes de conjugaison de $\mathfrak{A}_4$.

E34. Soit $n \geqslant 1$. On dit qu'un sous-groupe G de $\mathfrak{S}_n$ est *transitif* si l'action naturelle de G sur $[\![1, n]\!]$ est transitive.

a. Montrer que pour tout $n \geqslant 3$, $\mathfrak{A}_n$ est un sous-groupe transitif de $\mathfrak{S}_n$.

b. Soit G un sous-groupe transitif de $\mathfrak{S}_n$. Montrer que $n \mid |G|$. Donner un exemple d'un sous-groupe transitif de $\mathfrak{S}_n$ qui soit d'ordre n, pour tout $n \geqslant 1$.

c. Montrer que tout groupe d'ordre n est isomorphe à un sous-groupe transitif d'ordre n de $\mathfrak{S}_n$.

d. Réciproquement, tout sous-groupe d'ordre n de $\mathfrak{S}_n$ est-il transitif ?

e. Soit G un sous-groupe transitif de $\mathfrak{S}_n$, avec $n \geqslant 2$. Vérifier que pour tout $i \in [\![1, n]\!]$, $\mathrm{Stab}_G(i)$ est conjugué à $\mathrm{Stab}_G(1)$, et en déduire que le groupe G contient une permutation sans points fixes. On pourra utiliser les résultats de l'exercice VI-29.

Chapitre IX

Groupes quotients

1. Définition

Soit G un groupe, et soit H un sous-groupe. Rappelons que nous avons défini une relation d'équivalence sur G de la manière suivante. Si x, $y \in G$, alors $x \sim y$ si $x^{-1}y \in H$. Les classes d'équivalence pour cette relation sont les classes à gauche modulo H. Autrement dit, pour tout $x \in G$, on a

$$\overline{x} = xH = \{xh \mid h \in H\}.$$

On note alors G/H l'ensemble des classes d'équivalence. Les éléments de l'ensemble G/H sont donc des **parties** de G. Ce ne sont **pas** des éléments de G.

On a alors une application surjective

$$\begin{aligned} \pi : G &\longrightarrow G/H \\ x &\longmapsto \overline{x}. \end{aligned}$$

On voudrait définir une structure de groupe sur G/H de manière à ce que la projection canonique $\pi : G \longrightarrow G/H$ soit un morphisme de groupes. La seule façon pour cela serait de poser

$$\overline{x}\ \overline{y} = \overline{xy}, \qquad \text{pour tous } x, y \in G,$$

cela traduisant exactement ce que l'on veut. Malheureusement, rien ne dit que cette loi soit bien définie.

En effet, la difficulté cachée est que le résultat ne doit pas dépendre des éléments choisis pour représenter les classes d'équivalence de x et de y. Autrement dit, on doit vérifier que si $\overline{x}_1 = \overline{x}_2$ et $\overline{y}_1 = \overline{y}_2$, alors $\overline{x_1 y_1} = \overline{x_2 y_2}$. Ce n'est pas toujours le cas, comme le montre l'exemple suivant.

1.1. Exemple. Soit $G = \mathfrak{S}_3$ et $H = \langle (1\ 2) \rangle$. Soient

$$x_1 = (1\ 3), \quad x_2 = (1\ 3)(1\ 2) = (1\ 2\ 3), \quad y_1 = y_2 = (2\ 3).$$

Alors, $x_1 \sim x_2$, $y_1 \sim y_2$, et donc $\overline{x}_1 = \overline{x}_2$ et $\overline{y}_1 = \overline{y}_2$. En revanche, on a

$$x_1 y_1 = (1\ 3\ 2) \notin H, \quad x_2 y_2 = (1\ 2) \in H,$$

et donc $\overline{x_1 y_1} \neq \overline{x_2 y_2}$.

Examinons le problème d'un peu plus près. Supposons que la loi interne

$$\begin{aligned} G/H \times G/H &\longrightarrow G/H \\ (\overline{x}, \overline{y}) &\longmapsto \overline{xy} \end{aligned}$$

soit bien définie. Dans ce cas, on a

$$\overline{g} = \overline{1_G g} = \overline{1}_G\ \overline{g}, \quad \text{pour tout } g \in G.$$

Mais, pour tout $h \in H$, on a $\overline{1}_G = \overline{h}$ par définition de l'ensemble G/H, et ainsi, on obtient

$$\overline{g} = \overline{h}\ \overline{g} = \overline{hg}, \quad \text{pour tout } g \in G \text{ et tout } h \in H.$$

Par définition de la relation de congruence modulo H, cela revient à

$$g^{-1} h g \in H, \quad \text{pour tout } g \in G \text{ et tout } h \in H.$$

Puisque le passage à l'inverse induit une bijection de G sur lui-même, en remplaçant g par g^{-1} dans l'énoncé précédent, on constate que cela équivaut finalement à

$$ghg^{-1} \in H, \quad \text{pour tout } g \in G \text{ et tout } h \in H.$$

Autrement dit, H est nécessairement un sous-groupe distingué de G.

Cela explique d'ailleurs l'exemple précédent, puisque $\langle (1\ 2) \rangle$ n'est pas distingué dans $\mathfrak{S}_3$.

La proposition suivante montre que, réciproquement, si H est un sous-groupe distingué de G, on peut munir G/H d'une structure de groupe ayant les propriétés voulues.

1.2. Proposition. *Soit H un sous-groupe distingué de G. Alors, la loi interne*

$$\begin{aligned} G/H \times G/H &\longrightarrow G/H \\ (\overline{x}, \overline{y}) &\longmapsto \overline{xy} \end{aligned}$$

est bien définie, de neutre $\overline{1}_G$, et induit sur G/H une structure de groupe. De plus, l'application

$$\begin{aligned} \pi : G &\longrightarrow G/H \\ x &\longmapsto \overline{x} \end{aligned}$$

est un morphisme de groupes.

Démonstration. Vérifions que si $\overline{x}_1 = \overline{x}_2$ et $\overline{y}_1 = \overline{y}_2$, alors $\overline{x_1 y_1} = \overline{x_2 y_2}$. Par définition de la relation d'équivalence, il existe $h, h' \in H$ tels que

$$x_2 = x_1 h, \quad y_2 = y_1 h'.$$

On a alors

$$x_2 y_2 = x_1 h y_1 h' = x_1 y_1 (y_1^{-1} h y_1) h'.$$

Comme H est un sous-groupe distingué de G, on a $y_1^{-1} h y_1 \in H$, et par conséquent $(y_1^{-1} h y_1) h' \in H$. Ainsi, $\overline{x_1 y_1} = \overline{x_2 y_2}$, et l'application

$$\begin{aligned} G/H \times G/H &\longrightarrow G/H \\ (\overline{x}, \overline{y}) &\longmapsto \overline{xy} \end{aligned}$$

est bien définie.

Vérifions maintenant que G/H, muni de cette loi interne, est bien un groupe. Pour tout $\overline{x} \in G/H$, on a

$$\overline{1}_G \overline{x} = \overline{1_G x} = \overline{x},$$

et de même $\overline{x}\overline{1}_G = \overline{x}$. Donc, $\overline{1}_G$ est un élément neutre pour cette loi. De plus, la loi est associative. En effet, pour tous $\overline{x}$, $\overline{y}$, $\overline{z} \in G/H$, on a

$$(\overline{x}\ \overline{y})\overline{z} = \overline{xy}\ \overline{z} = \overline{(xy)z} = \overline{x(yz)} = \overline{x}\ \overline{yz} = \overline{x}(\overline{y}\ \overline{z}).$$

Enfin, on vérifie facilement que tout $\overline{x} \in G/H$ admet pour inverse la classe à gauche $\overline{x^{-1}}$.

Le fait que π soit un morphisme de groupes provient de la définition de la loi de groupe sur G/H. □

1.3. Définition. Soit G un groupe, et soit H un sous-groupe distingué dans G. Le groupe G/H est appelé le *groupe quotient* de G par H.

1.4. Remarques

(1) La chose la plus importante à retenir de cette définition, outre la définition de la loi de groupe, est la suivante : pour tout $x \in G$, on a

$$\overline{x} = \overline{1}_G \iff x \in H.$$

(2) Si $[G : H]$ est fini, G/H est d'ordre $[G : H]$, par définition. En particulier, si G est fini, par le théorème de Lagrange, on a

$$|G/H| = \frac{|G|}{|H|}.$$

(3) Si G est abélien, tout sous-groupe H est distingué dans G, et l'on peut toujours former le groupe quotient G/H.

(4) La projection canonique $\pi : G \longrightarrow G/H$ est surjective, de noyau H. Ainsi, tout sous-groupe distingué de G est le noyau d'un certain morphisme de groupes.

Quel est l'intérêt d'un groupe quotient ? En fait, quotienter G par un sous-groupe H distingué dans G revient à imposer des relations entre les éléments du groupe G qui n'existent pas forcément. En un sens, « on force les éléments de H à devenir triviaux ».

1.5. Exemples

(1) On a $G/G = \{\overline{1}_G\}$. En effet, tous les éléments sont équivalents dans ce cas. En fait, le groupe G/H est trivial si, et seulement si, $H = G$.

(2) On a $G \simeq G/\{1_G\}$, l'isomorphisme étant donné par la projection canonique. En effet, on a $\overline{x} = \{x\}$, pour tout $x \in G$. La projection canonique est donc bien un isomorphisme. Attention, on a un isomorphisme, mais **pas** une égalité !!

(3) Soit $n \geqslant 1$. Le groupe $\mathbb{Z}/n\mathbb{Z}$ n'est rien d'autre que le quotient de $\mathbb{Z}$ par le sous-groupe $n\mathbb{Z}$.

On finit ce paragraphe en étudiant les sous-groupes de G/H.

1.6. Proposition. *Soit G un groupe et soit H un sous-groupe distingué de G. Alors, la projection canonique $\pi : G \longrightarrow G/H$ induit une correspondance bijective entre l'ensemble des sous-groupes H' de G contenant H et l'ensemble des sous-groupes de G/H.*

Plus précisément, si H' est un sous-groupe de G contenant H, alors

$$H'/H = \pi(H') = \left\{\overline{h'} \mid h' \in H'\right\}$$

est un sous-groupe de G/H.

Réciproquement, si S est un sous-groupe de G/H, alors $\pi^{-1}(S)$ est un sous-groupe de G contenant H, et ces deux constructions sont inverses l'une de l'autre.

De plus, cette correspondance établit une bijection entre l'ensemble des sous-groupes distingués de G contenant H et l'ensemble des sous-groupes distingués de G/H.

Démonstration. Puisque π est un morphisme de groupes, $\pi(H')$ et $\pi^{-1}(S)$ sont des sous-groupes de G/H et G respectivement par le lemme VI-2.5. De plus, pour tout $h \in H$, on a $\pi(h) = \overline{h} = \overline{1}_G \in S$, et donc $H \subset \pi^{-1}(S)$. Il reste à vérifier que les deux constructions sont inverses l'une de l'autre. Soit H' un sous-groupe de G contenant H, et soit $S = \pi(H') = H'/H$.

Alors, on a

$$\pi^{-1}(S) = \{x \in G \mid \overline{x} \in H'/H\}.$$

Il reste à constater que $\{x \in G \mid \overline{x} \in H'/H\} = H'$.

L'inclusion « $\supset$ » est claire. Si maintenant $\overline{x} \in H'/H$, il existe $y \in H'$ tel que $\overline{x} = \overline{y}$. On a donc $x = yh, h \in H$. Or, $H \subset H'$, et ainsi $x \in H'$. On a donc bien $\pi^{-1}(S) = H'$.

Inversement, soit S un sous-groupe de G/H, et soit $H' = \pi^{-1}(S)$. On a donc $\pi(H') \subset S$ par définition de H'. Soit $\overline{x} \in S$. Alors, $\overline{x} = \pi(x) \in S$, et donc $x \in H'$. Ainsi, on a $\overline{x} \in \pi(H')$, et donc $S = \pi(H')$.

Si H' est un sous-groupe distingué de G contenant H, alors H'/H est distingué dans G/H. En effet, pour tout $\overline{x} \in G/H$ et tout $h' \in H'$, on a

$$\overline{x}\overline{h'}\overline{x}^{-1} = \overline{xh'x^{-1}} \in H'/H,$$

car H' est distingué dans H. Enfin, soit S un sous-groupe distingué de G/H. Alors, pour tout $x \in G$ et tout $h \in \pi^{-1}(S)$, on a

$$\pi(xhx^{-1}) = \pi(x)\pi(h)\pi(x)^{-1} \in S,$$

car $\pi(h) \in S$ et S est distingué dans G/H. Ainsi, $\pi^{-1}(S)$ est distingué dans G. Cela achève la démonstration. □

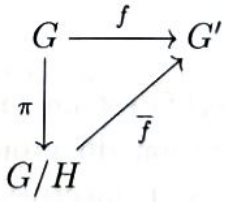
2. Théorème de factorisation

Soit G un groupe, et soit H un sous-groupe distingué de G. Nous allons dans ce paragraphe nous intéresser aux morphismes de G/H dans un groupe G'.

2.1. Théorème (de factorisation). *Soit G un groupe, et soit H un sous-groupe distingué de G. Soit $f : G \longrightarrow G'$ un morphisme de groupes tel que $H \subset \mathrm{Ker}(f)$. Alors, il existe un unique morphisme de groupes*

$$\overline{f} : G/H \longrightarrow G'$$

tel que $f = \overline{f} \circ \pi$.



Ce morphisme de groupes est défini par

$$\overline{f}(\overline{x}) = f(x), \quad \text{pour tout } \overline{x} \in G/H.$$

De plus, pour tout groupe G', il y a une correspondance bijective

$$\begin{array}{ccc} \mathrm{Hom}(G/H, G') & \longrightarrow & \{f \in \mathrm{Hom}(G, G') \mid H \subset \mathrm{Ker}(f)\} \\ \varphi & \longmapsto & \varphi \circ \pi \\ \overline{f} & \longleftarrow & f. \end{array}$$

Démonstration. Remarquons tout d'abord qu'un morphisme $f : G \longrightarrow G'$ est constant sur les classes d'équivalence (pour la relation associée à H) si, et seulement si, $H \subset \mathrm{Ker}(f)$.

En effet, si $H \subset \mathrm{Ker}(f)$, alors pour tout $x \in G$ et $h \in H$, on a

$$f(xh) = f(x)f(h) = f(x).$$

Réciproquement, si f est constante sur chaque classe d'équivalence, elle est constante sur H. Mais H contient 1, et l'on a donc

$$f(h) = f(1_G) = 1_{G'}, \quad \text{pour tout } h \in H,$$

c'est-à-dire $H \subset \mathrm{Ker}(f)$.

La proposition I-2.11 fournit alors l'existence et l'unicité d'une application $\overline{f}$ vérifiant les propriétés voulues. De plus, pour tous x_1, $x_2 \in G$, on a

$$\overline{f}(\overline{x}_1\overline{x}_2) = \overline{f}(\overline{x_1x_2}) = f(x_1x_2) = f(x_1)f(x_2) = \overline{f}(\overline{x}_1)\overline{f}(\overline{x}_2),$$

et $\overline{f}$ est donc un morphisme de groupes.

Démontrons la seconde partie. Par la proposition I-2.11, on a une bijection entre l'ensemble des applications de G/H dans G' et l'ensemble des applications $G \longrightarrow G'$ constantes sur les classes d'équivalence. Elle est donnée par

$$\begin{array}{ccc} \varphi & \longmapsto & \varphi \circ \pi \\ \overline{f} & \longmapsfrom & f\,. \end{array}$$

Si maintenant φ est un morphisme de groupes, alors $\varphi \circ \pi$ est un morphisme dont le noyau contient H. De plus, si f est un morphisme de groupes, alors $\overline{f}$ est un morphisme de groupes.

Ainsi, la correspondance se restreint pour définir une bijection entre les ensembles voulus. Cela achève la démonstration. □

Remarquons en particulier que tout morphisme de groupes $f : G \longrightarrow G'$ induit un morphisme de groupes

$$\overline{f} : G/\mathrm{Ker}(f) \longrightarrow G',$$

puisque $\mathrm{Ker}(f)$ est un sous-groupe distingué de G qui contient... $\mathrm{Ker}(f)$. Vu la définition du groupe quotient, on peut s'attendre à ce que cette application soit injective, puisque l'on « tue » tous les éléments du noyau. C'est en effet le cas.

2.2. Lemme. *Soit $f : G \longrightarrow G'$ un morphisme de groupes. Alors, le morphisme de groupes*

$$\overline{f} : G/\mathrm{Ker}(f) \longrightarrow G'$$

est injectif.

Démonstration. Soit $\overline{x} \in G/\mathrm{Ker}(f)$. On a $\overline{f}(\overline{x}) = 1_{G'}$ si, et seulement si, on a $f(x) = 1_{G'}$, c'est-à-dire si, et seulement si, $x \in \mathrm{Ker}(f)$, ce qui revient

encore à dire que $\overline{x} = \overline{1}_G \in G/\mathrm{Ker}(f)$. Ainsi, $\mathrm{Ker}(\overline{f})$ est réduit à l'élément neutre et $\overline{f}$ est injective. □

On obtient alors le résultat fondamental suivant.

2.3. Théorème (Premier théorème d'isomorphisme). *Soient G, G' deux groupes, et soit $f : G \longrightarrow G'$ un morphisme de groupes. Alors, le morphisme de groupes*

$$\overline{f} : G/\mathrm{Ker}(f) \longrightarrow G'$$

induit un isomorphisme de groupes

$$G/\mathrm{Ker}(f) \simeq \mathrm{Im}(f).$$

Démonstration. D'après le lemme précédent, $\overline{f} : G/\mathrm{Ker}(f) \longrightarrow G'$ est un morphisme injectif, et induit donc un isomorphisme de $G/\mathrm{Ker}(f)$ sur son image. Or, on a

$$\mathrm{Im}(\overline{f}) = \left\{\overline{f}(\overline{x}) \mid \overline{x} \in G/\mathrm{Ker}(f)\right\} = \left\{f(x) \mid x \in G\right\} = \mathrm{Im}(f),$$

d'où le résultat. □

Le théorème 2.3 nous permet en pratique d'identifier un groupe quotient G/H à un groupe connu de la manière suivante : on essaye d'identifier H au noyau d'un morphisme de groupes $f : G \longrightarrow G'$ bien choisi, et l'on calcule l'image de f (en pratique, on essaye de trouver un f qui soit surjectif).

2.4. Exemples

(1) Identifions le quotient $\mathbb{C}^\times/\mathbb{U}$, où $\mathbb{U} = \{z \in \mathbb{C}^\times \mid |z| = 1\}$, et considérons l'application

$$\begin{aligned} f : \mathbb{C}^\times &\longrightarrow \mathbb{R}^{+\times} \\ z &\longmapsto |z|. \end{aligned}$$

C'est un morphisme de groupes surjectif, de noyau $\mathbb{U}$, et le premier théorème d'isomorphisme nous donne

$$\mathbb{C}^\times/\mathbb{U} \simeq \mathbb{R}^{+\times}.$$

(2) Identifions $\mathbb{R}/2\pi\mathbb{Z}$. On considère le morphisme de groupes surjectif

$$\begin{aligned} f : \mathbb{R} &\longrightarrow \mathbb{U} \\ \theta &\longmapsto e^{i\theta}. \end{aligned}$$

Son noyau étant égal à $2\pi\mathbb{Z}$, on obtient

$$\mathbb{R}/2\pi\mathbb{Z} \simeq \mathbb{U}.$$

Nous allons maintenant démontrer deux autres théorèmes d'isomorphisme.

2.5. Théorème (Deuxième théorème d'isomorphisme). *Soit G un groupe, et soient H et K deux sous-groupes. On suppose que H est distingué dans G. Alors, on a les propriétés suivantes :*

(1) *les ensembles*

$$HK = \{hk \mid h \in H, k \in K\} \quad \text{et} \quad KH = \{kh \mid h \in H, k \in K\}$$

sont des sous-groupes de G. De plus, $\langle H, K\rangle = HK = KH$;

(2) *le sous-groupe H est distingué dans HK, et $H \cap K$ est distingué dans K, de sorte que l'on a un isomorphisme de groupes*

$$HK/H \simeq K/(H \cap K).$$

Démonstration

(1) Montrons que HK et KH sont des sous-groupes de G.

Puisque H et K sont des sous-groupes de G, ils contiennent 1_G, et donc HK contient $1_G 1_G = 1_G$. Si $h_1k_1, h_2k_2 \in HK$, on a

$$(h_1k_1)(h_2k_2) = \big(h_1(k_1h_2k_1^{-1})\big)(k_1k_2) \in HK,$$

puisque H est distingué dans G, et H et K sont des sous-groupes. Enfin, si $hk \in HK$, on a également

$$(hk)^{-1} = k^{-1}h^{-1} = (k^{-1}h^{-1}k)k^{-1} \in HK,$$

pour les mêmes raisons que précédemment. Ainsi, HK est bien un sous-groupe de G.

Montrons maintenant l'égalité $\langle H, K\rangle = HK$.

Puisque H et K sont contenus dans $\langle H, K\rangle$, on a $HK \subset \langle H, K\rangle$. Comme HK est un sous-groupe de G contenant H et K, on a $\langle H, K\rangle \subset HK$, d'où le résultat.

Il reste à voir que $HK = KH$, ce qui montrera toutes les assertions manquantes. Comme H, K et HK sont des sous-groupes de G, on a

$$HK = (HK)^{-1} = K^{-1}H^{-1} = KH,$$

où pour toute partie P de G, P^{-1} désigne l'ensemble des inverses des éléments de P. Cela achève la démonstration de (1).

(2) Pour tout $h' \in H$, et tout $hk \in HK$, on a

$$(hk)h'(hk)^{-1} = h(kh'k^{-1})h^{-1} \in H,$$

et donc H est distingué dans HK.

On peut ainsi former le groupe quotient HK/H. Considérons le morphisme de groupes

$$\begin{aligned} f : K &\longrightarrow HK/H \\ k &\longmapsto \overline{k}. \end{aligned}$$

Pour tout $hk \in HK$, on a

$$\overline{hk} = \overline{h}\,\overline{k} = \overline{1}\,\overline{k} = \overline{k} = f(k),$$

et ainsi f est surjective. Par ailleurs, pour tout $k \in K$, on a

$$\overline{k} = \overline{1} \iff k \in H.$$

Autrement dit, $\mathrm{Ker}(f) = H \cap K$. Ainsi, $H \cap K$ est distingué dans K par le lemme VI-2.15, et par le premier théorème d'isomorphisme, on obtient

$$K/(H \cap K) \simeq HK/H,$$

d'où le résultat. □

2.6. Théorème (Troisième théorème d'isomorphisme). *Soit G un groupe, et soient H et K deux sous-groupes. On suppose que H et K sont distingués dans G, et que K est un sous-groupe de H. Alors, K est distingué dans H, H/K est distingué dans G/K, et l'on a un isomorphisme de groupes*

$$(G/K)/(H/K) \simeq G/H.$$

Démonstration. Clairement, K est distingué dans H. Soit $\pi : G \longrightarrow G/H$ la projection canonique. Comme $K \subset H = \mathrm{Ker}(\pi)$, le théorème de factorisation montre que π induit un morphisme de groupes

$$\begin{aligned} \tilde{\pi} : G/K &\longrightarrow G/H \\ \tilde{x} &\longmapsto \overline{x}, \end{aligned}$$

où $\tilde{x}$ et $\overline{x}$ désignent respectivement les classes à gauche de x modulo K et modulo H.

Le morphisme $\tilde{\pi}$ est évidemment surjectif. De plus, pour tout $x \in G$, on a

$$\overline{x} = \overline{1} \iff x \in H \iff \tilde{x} \in H/K.$$

Autrement dit, $\mathrm{Ker}(\tilde{\pi}) = H/K$. Ainsi, H/K est distingué dans G/K, et par le premier d'isomorphisme, on a

$$(G/K)/(H/K) \simeq G/H,$$

ce qui achève la démonstration. □

Nous finissons ce paragraphe par une jolie application du théorème de factorisation.

2.7. Théorème. *Pour tous groupes* **finis** G, G_1, G_2, *on a*

$$G_1 \times G \simeq G_2 \times G \quad \Longrightarrow \quad G_1 \simeq G_2.$$

Démonstration. Si K et K' sont deux groupes, on note $\mathrm{Hom}(K, K')$ l'ensemble des morphismes de groupes $K \longrightarrow K'$, et $\mathrm{Hom}^0(K, K')$ l'ensemble des morphismes de groupes **injectifs** $K \longrightarrow K'$. Quand ces ensembles sont finis, on note $h(K, K')$ et $h^0(K, K')$ leurs nombres d'éléments respectifs.

Avant d'aborder la démonstration proprement dite, faisons quelques remarques préliminaires.

(1) Tout d'abord, pour tous groupes K, K'_1, K'_2, on a une bijection entre les ensembles $\mathrm{Hom}(K, K'_1 \times K'_2)$ et $\mathrm{Hom}(K, K'_1) \times \mathrm{Hom}(K, K'_2)$.

En effet, notons $\pi_1 : K'_1 \times K'_2 \longrightarrow K'_1$ et $\pi_2 : K'_1 \times K'_2 \longrightarrow K'_2$ les deux projections. Ce sont bien entendu des morphismes de groupes. Mais, alors, pour tout $f \in \mathrm{Hom}(K, K'_1 \times K'_2)$, on a $\pi_i \circ f \in \mathrm{Hom}(K, K'_i)$ pour $i = 1, 2$. On a donc une application

$$\begin{array}{ccc} \mathrm{Hom}(K, K'_1 \times K'_2) & \longrightarrow & \mathrm{Hom}(K, K'_1) \times \mathrm{Hom}(K, K'_2) \\ f & \longmapsto & (\pi_1 \circ f, \pi_2, \circ f), \end{array}$$

qui est bijective, puisque une application $K \longrightarrow K'_1 \times K'_2$ est déterminée de manière unique par ses projections sur K'_1 et K'_2.

En particulier, si K, K'_1, K'_2 sont des groupes finis, on a

$$h(K, K'_1 \times K'_2) = h(K, K'_1) h(K, K'_2).$$

(2) Soient K et K' deux groupes. Alors, on a une bijection entre les ensembles $\mathrm{Hom}(K, K')$ et $E = \bigcup_{H \lhd K} \mathrm{Hom}^0(K/H, K')$.

Notons que cette dernière réunion est disjointe. En effet, la seule classe à gauche de K modulo H contenant 1_K est H. En particulier, si H_1, H_2 sont deux sous-groupes distingués de K tels que $K/H_1 = K/H_2$, alors $H_1 = H_2$.

Établissons la bijection désirée. Rappelons tout d'abord que pour tout morphisme de groupes $f : K \longrightarrow K'$, le morphisme $\overline{f} : K/\mathrm{Ker}(f) \longrightarrow K'$ est injectif (cf. lemme 2.2). On a donc une application

$$\begin{array}{ccc} \theta : \mathrm{Hom}(K, K') & \longrightarrow & E \\ f & \longmapsto & \overline{f}. \end{array}$$

Si f, $g \in \mathrm{Hom}(K, K')$ sont deux morphismes de groupes tels que $\overline{f} = \overline{g}$, alors on a $K/\mathrm{Ker}(f) = K/\mathrm{Ker}(g)$, et en composant à droite par la projection canonique $K \longrightarrow K/\mathrm{Ker}(f)$, on obtient $f = g$. Ainsi, θ est injective.

Montrons maintenant que θ est surjective. Soit $\varphi : K/H \longrightarrow K'$ est un morphisme de groupes injectifs, où H est un sous-groupe distingué de K.

Soit $\pi : K \longrightarrow K/H$ la projection canonique. Alors, pour tout $x \in K$, on a

$$(\varphi \circ \pi)(x) = 1_{K'} \iff \varphi(\overline{x}) = 1_{K'} \iff \overline{x'} = \overline{1}_K \iff x \in H.$$

Autrement dit, $\mathrm{Ker}(\varphi \circ \pi) = H$. Or, la dernière partie du théorème de factorisation affirme le morphisme φ est obtenu de $\varphi \circ \pi$ par passage au quotient. On en déduit que $\theta(\varphi \circ \pi) = \varphi$, d'où la surjectivité de θ.

On a donc bien la bijection annoncée. En particulier, si K et K' sont finis, on a

$$h(K, K') = \sum_{H \lhd K} h^0(K/H, K').$$

Passons maintenant à la démonstration proprement dite. Soient G, G_1, G_2 des groupes finis tels que $G_1 \times G \simeq G_2 \times G$. Nous allons démontrer que pour tout groupe fini K, on a $h^0(K, G_1) = h^0(K, G_2)$.

Pour tout entier $n \geqslant 1$, soit (H_n) la propriété :

(H_n) Pour tout groupe fini K tel que $|K| \leqslant n$, on a $h^0(K, G_1) = h^0(K, G_2)$.

Montrons que (H_n) est vraie pour tout $n \geqslant 1$ par récurrence. Si $n = 1$, K est trivial, et le résultat est clair. Supposons que (H_n) soit vraie pour un certain entier $n \geqslant 1$, et montrons que (H_{n+1}) est vraie. Soit K un groupe fini tel que $|K| \leqslant n+1$. Puisque $G_1 \times G \simeq G_2 \times G$, on a une bijection entre les ensembles $\mathrm{Hom}(K, G_1 \times G)$ et $\mathrm{Hom}(K, G_2 \times G)$, et ainsi

$$h(K, G_1 \times G) = h(K, G_2 \times G).$$

D'après le point (1), on a donc $h(K, G_1)h(K, G) = h(K, G_2)h(K, G)$. Puisque l'entier $h(K, G)$ est non nul (il existe toujours au moins un morphisme de groupes $K \longrightarrow G$, à savoir le morphisme trivial), on en déduit l'égalité

$$h(K, G_1) = h(K, G_2).$$

D'après le point (2), on a alors

$$\sum_{H \lhd K} h^0(K/H, G_1) = \sum_{H \lhd K} h^0(K/H, G_2).$$

Si H est un sous-groupe distingué de K non trivial, on a $|K/H| < |K|$, et donc $|K/H| \leqslant n$. Par hypothèse de récurrence, on obtient

$$h^0(K/H, G_1) = h^0(K/H, G_2),$$

et l'on en déduit, en simplifiant les termes dans l'égalité précédente, que

$$h^0(K, G_1) = h^0(K, G_2).$$

Cela achève la récurrence.

Nous pouvons maintenant démontrer le résultat attendu. Puisque Id_{G_1} est

un morphisme de groupes injectif, le point précédent montre que l'on a

$$h^0(G_1, G_2) = h^0(G_1, G_1) \geqslant 1.$$

Ainsi, il existe au moins un morphisme de groupes injectif $G_1 \longrightarrow G_2$. Mais, l'isomorphisme $G_1 \times G \simeq G_2 \times G$ implique en particulier que les groupes G_1 et G_2 ont même ordre. Le morphisme précédent est donc un isomorphisme, d'où le résultat souhaité [1]. □

2.8. Remarque. Ce théorème devient faux pour les groupes infinis. En effet, il est facile de vérifier que l'application

$$\begin{aligned} \mathbb{Z}^{(\mathbb{N})} &\longrightarrow \quad \mathbb{Z} \times \mathbb{Z}^{(\mathbb{N})} \\ (u_n)_{n\geqslant 0} &\longmapsto \big(u_0, (u_{n+1})_{n\geqslant 0}\big) \end{aligned}$$

est un isomorphisme de groupes. On a ainsi $\{0\} \times \mathbb{Z}^{(\mathbb{N})} \simeq \mathbb{Z} \times \mathbb{Z}^{(\mathbb{N})}$, bien que le groupe $\mathbb{Z}$ soit non trivial. En revanche, le théorème reste vrai pour les groupes abéliens de type fini (cf. exercice XII-10).

3. Suites de composition

Dans ce paragraphe, nous allons très brièvement montrer comment la notion de groupe quotient permet de « dévisser » la structure d'un groupe. Nous obtiendrons ainsi un analogue de la décomposition en facteurs premiers pour les groupes.

Commençons par définir la notion de suite de composition.

3.1. Définition. Soit G un groupe. Une *suite de composition* de G est une suite finie $(G_0, \ldots, G_r)$ de sous-groupes de G, telle que $G_0 = G$, $G_r = \{1_G\}$ et pour tout $i \in [\![0, r-1]\!]$, G_{i+1} est un sous-groupe distingué dans G_i.

On a donc

$$\{1_G\} = G_r \lhd G_{r-1} \lhd \cdots \lhd G_1 \lhd G_0 = G.$$

Les groupes G_i/G_{i+1} sont appelés les *quotients* de la suite de composition, et l'entier r est appelé la *longueur* de la suite de composition.

3.2. Exemples

(1) Pour tout groupe G, la suite $(G, \{1_G\})$ est une suite de composition.

(2) Un groupe simple G n'admet que des suites de composition triviales, de la forme $(G, \ldots, G, \{1_G\}, \ldots, \{1_G\})$.

1. Les arguments présentés ici sont une version détaillée de ceux décrits dans http ://www.les-mathematiques.net/phorum/read.php ?3,1463332,1835882.

(3) Pour tout sous-groupe distingué H de G, la suite $(G, H, \{1_G\})$ est une suite de composition.

Le lemme technique suivant explique comment fabriquer des suites de composition de sous-groupes et de groupes quotients à partir d'une suite de composition de G, et inversement.

3.3. Lemme. *Soit G un groupe.*

(1) *Soit H un sous-groupe de G. Si $(G_0, \ldots, G_r)$ est une suite de composition de G, la suite*
$$(G_0 \cap H, \ \ldots, \ G_r \cap H)$$
est une suite de composition de H.
De plus, pour tout $i \in [\![0, r-1]\!]$, $(G_i \cap H)/(G_{i+1} \cap H)$ est isomorphe à un sous-groupe de G_i/G_{i+1}.

(2) *Soit H un sous-groupe distingué de G. Si $(G_0, \ldots, G_r)$ est une suite de composition de G, la suite*
$$(\langle G_0, H\rangle/H, \ \ldots, \ \langle G_r, H\rangle/H)$$
est une suite de composition de G/H.
De plus, pour tout $i \in [\![0, r-1]\!]$, $(\langle G_i, H\rangle/H)/(\langle G_{i+1}, H\rangle/H)$ est isomorphe à un quotient de G_i/G_{i+1}.

(3) *Soit H un sous-groupe distingué de G. Si $(H_0, \ldots, H_r)$ est une suite de composition de H, et si $(G'_0/H, \ldots, G'_s/H)$ est une suite de composition du quotient G/H (où G'_i est un sous-groupe de G contenant H), alors la suite*
$$(G'_0, \ \ldots, \ G'_s, \ H_1, \ \ldots, \ H_r)$$
est une suite de composition de G, dont la liste des quotients successifs s'obtient en concaténant les listes des quotients successifs des deux suites ci-dessus.

Démonstration

(1) Puisque $G_0 = G$ et $G_r = \{1_G\}$, on a $G_0 \cap H = H$, ainsi que
$$G_r \cap H = \{1_G\} = \{1_H\}.$$
De plus, pour tout $i \in [\![0, r-1]\!]$, $G_{i+1} \cap H$ est distingué dans $G_i \cap H$, car G_{i+1} est un sous-groupe distingué de G_i.
Remarquons maintenant que le morphisme de groupes
$$\begin{aligned} G_i \cap H &\longrightarrow G_i/G_{i+1} \\ g_i &\longmapsto \overline{g}_i \end{aligned}$$
est de noyau $(G_i \cap H) \cap G_{i+1} = G_{i+1} \cap H$, et induit donc par passage

au quotient un morphisme de groupes injectif

$$\begin{array}{rccc} \varphi_i : & (G_i \cap H)/(G_{i+1} \cap H) & \longrightarrow & G_i/G_{i+1} \\ & \tilde{g}_i & \longmapsto & \overline{g}_i, \end{array}$$

d'après le lemme 2.2.

(2) Puisque $G_0 = G$ et $G_r = \{1_G\}$, on a $\langle G_0, H\rangle = G$ et $\langle G_r, H\rangle = H$. De plus, pour tout $i \in [\![0, r-1]\!]$, $\langle G_{i+1}, H\rangle$ est distingué dans $\langle G_i, H\rangle$.

En effet, pour tous h, $h' \in H$, tout $g_i \in G_i$, et tout $g_{i+1} \in G_{i+1}$, on a

$$g_i h' g_i^{-1} \in \langle G_i, H\rangle$$

par définition d'un sous-groupe engendré (ou bien parce que H est distingué dans G), et $hh'h^{-1} \in H$. Ainsi, H est distingué dans $\langle G_i, H\rangle$. De plus, on a

$$hg_{i+1}h^{-1} \in \langle G_{i+1}, H\rangle \subset \langle G_i, H\rangle,$$

et aussi

$$g_i g_{i+1} g_i^{-1} \in G_{i+1} \subset \langle G_i, H\rangle.$$

Cela montre que G_{i+1} est aussi distingué dans $\langle G_i, H\rangle$. Par conséquent, le sous-groupe $\langle G_{i+1}, H\rangle$ est bien distingué dans $\langle G_i, H\rangle$.

Considérons le morphisme de groupes

$$\begin{array}{rccc} f_i : & G_i & \longrightarrow & \langle G_i, H\rangle/\langle G_{i+1}, H\rangle \\ & g_i & \longmapsto & \overline{g}_i. \end{array}$$

Alors, f_i est surjectif.

En effet, par le théorème 2.5, pour tout $i \in [\![0, r-1]\!]$, on a

$$\langle G_i, H\rangle = \langle H, G_i\rangle = HG_i,$$

car H est distingué dans G. Pour tout $hg_i \in HG_i = \langle H, G_i\rangle$, on a alors

$$\overline{hg_i} = \overline{h}\ \overline{g}_i = \overline{g}_i = f_i(g_i),$$

d'où la surjectivité [2].

De plus, $\mathrm{Ker}(f_i)$ contient G_{i+1}, et induit donc par passage au quotient un morphisme de groupes

$$\overline{f}_i : G_i/G_{i+1} \longrightarrow \langle G_i, H\rangle/\langle G_{i+1}, H\rangle.$$

Ce morphisme est surjectif, puisque f_i l'est. Le premier théorème d'isomorphisme permet alors de conclure que $\langle G_i, H\rangle/\langle G_{i+1}, H\rangle$ est isomorphe au quotient de G_i/G_{i+1} par $\mathrm{Ker}(f_i)$.

Enfin, par le troisième théorème d'isomorphisme, puisque H est distingué dans G, donc dans $\langle G_i, H\rangle$ et dans $\langle G_{i+1}, H\rangle$, on obtient l'iso-

2. On aurait aussi pu effectuer un raisonnement direct.

morphisme

$$\langle G_i, H\rangle/\langle G_{i+1}, H\rangle \simeq \big(\langle G_i, H\rangle/H\big)/\big(\langle G_{i+1}, H\rangle/H\big).$$

Ce dernier groupe est donc bien isomorphe à un quotient de G_i/G_{i+1}.

(3) Par définition, on a $H_r = \{1_H\} = \{1_G\}$, et $G'_0/H = G/H$. On a donc $G'_0 = G$. Remarquons également que le sous-groupe G'_{j+1} est distingué dans G'_j pour tout $j \in [\![0, s-1]\!]$ par la proposition 1.6. Enfin, puisque $G'_s/H = \{\overline{1}_G\}$, on a $G'_s = H$. Mais alors, H_1 est bien distingué dans $G'_s = H = H_0$. Ainsi, la suite

$$(G'_0, \ldots, G'_s, H_1, \ldots, H_r)$$

est une suite de composition de G. Le dernier point sur les quotients est clair par construction (en utilisant le fait que $G'_s = H = H_0$). □

Cette notion n'a évidemment d'intérêt que si l'on exige des propriétés supplémentaires sur les sous-groupes G_i ou sur les quotients de la suite de composition. Cela conduit à diverses notions importantes de théorie des groupes. Nous allons commencer par la notion de suite de Jordan-Hölder. Pour cela, introduisons une relation d'ordre sur l'ensemble des suites de composition.

3.4. Définition. Soient $\mathscr{G} = (G_0, \ldots, G_r)$ et $\mathscr{G}' = (G'_0, \ldots, G'_s)$ deux suites de composition d'un même groupe G.

On dit que $\mathscr{G}'$ est un *raffinement* de $\mathscr{G}$ si $\mathscr{G}$ est une suite extraite de $\mathscr{G}'$, c'est-à-dire s'il existe des indices $j_1, \ldots, j_r \in [\![0, s]\!]$ tels que

$$G'_{j_i} = G_i, \quad \text{pour tout } i \in [\![0, r]\!].$$

En particulier, $s \geqslant r$.

On dit aussi que $\mathscr{G}'$ est *plus fine* que $\mathscr{G}$, ou que $\mathscr{G}$ est *moins fine* que $\mathscr{G}'$. (On vérifie alors sans mal que la relation « être moins fine » est une relation d'ordre sur l'ensemble des suites de composition de G.)

Un *raffinement strict* de $\mathscr{G}$ est un raffinement $\mathscr{G}'$ distinct de $\mathscr{G}$. (Au vu de la définition, cela équivaut à $s > r$, avec les notations précédentes.)

Toute suite de composition $(G_0, \ldots, G_r)$ de G admet bien entendu un raffinement strict, par exemple la suite de composition $(G_0, G_0, G_1, \ldots, G_r)$. Pour que la relation d'ordre induite par le raffinement soit intéressante, il faut donc la restreindre à l'ensemble des suites de composition strictement décroissantes (ce qui est possible puisqu'un raffinement d'une suite de composition strictement décroissante est encore une suite de composition strictement décroissante). Le lemme suivant caractérise les éléments maximaux pour cette relation d'ordre.

3.5. Lemme. *Soit G un groupe, et soit $\mathscr{G} = (G_0, \ldots, G_r)$ une suite de composition de G. Les propriétés suivantes sont équivalentes :*

(1) *la suite $\mathscr{G}$ est strictement décroissante et ne possède pas de raffinement strict, i.e. $\mathscr{G}$ est un élément maximal de l'ensemble des suites de composition strictement décroissantes de G pour la relation d'ordre « être moins fine » ;*

(2) *pour tout $i \in [\![0, r-1]\!]$, le quotient G_i/G_{i+1} est un groupe simple.*

Démonstration. Notons que si $r = 0$, les conditions (1) et (2) sont automatiquement réalisées[3]. On peut donc supposer que $r \geqslant 1$.

Supposons tout d'abord que la suite $\mathscr{G}$ soit strictement décroissante et ne possède pas de raffinement strict et soit $i \in [\![0, r-1]\!]$. Puisque $\mathscr{G}$ est strictement décroissante, G_i/G_{i+1} est non trivial. De plus, un sous-groupe distingué de G_i/G_{i+1} est de la forme H_i/G_{i+1}, où H_i est un sous-groupe distingué de G_i contenant G_{i+1}, par la proposition 1.6. De plus, G_{i+1} étant distingué dans G_i, il l'est dans H_i. Ainsi, la suite de sous-groupes

$$(G_0, \ldots, G_i, H_i, G_{i+1}, \ldots, G_r)$$

est un raffinement strict de $\mathscr{G}$. Par hypothèse, cette suite de composition ne peut être strictement décroissante. Comme $\mathscr{G}$ est elle-même strictement décroissante, cela force à avoir $H_i = G_{i+1}$ ou $H_i = G_i$, i.e. H_i/G_{i+1} est soit trivial, soit égal à G_i/G_{i+1}. Ainsi, G_i/G_{i+1} est un groupe simple.

Supposons maintenant que pour tout $i \in [\![0, r-1]\!]$, le quotient G_i/G_{i+1} soit un groupe simple. Un groupe simple étant non trivial, on obtient

$$G_{i+1} \subsetneq G_i, \quad \text{pour tout } i \in [\![0, r-1]\!].$$

Ainsi, $\mathscr{G}$ est strictement décroissante. Soit $\mathscr{G}' = (G'_0, \ldots, G'_s)$ une suite de composition de G strictement décroissante qui soit un raffinement de $\mathscr{G}$. Alors, il existe des indices $j_1, \ldots, j_r \in [\![0, s]\!]$ tels que

$$G'_{j_i} = G_i, \quad \text{pour tout } i \in [\![0, r]\!].$$

Soit $i \in [\![0, r-1]\!]$. Le quotient $G'_{j_i}/G'_{j_{i+1}} = G_i/G_{i+1}$ est alors simple. Or, $G'_{j_i+1}/G'_{j_{i+1}}$ est distingué dans $G'_{j_i}/G'_{j_{i+1}}$. Il est donc soit, trivial, soit égal au groupe tout entier. On a donc $G'_{j_i+1} = G'_{j_{i+1}}$ ou $G'_{j_i+1} = G'_{j_i}$. Le second cas est impossible, car $\mathscr{G}'$ est strictement décroissante. La stricte décroissance de $\mathscr{G}'$ entraîne alors que

$$j_{i+1} = j_i + 1, \quad \text{pour tout } i \in [\![0, r-1]\!].$$

Sachant que $G'_{j_0} = G = G'_0$ et que la suite $\mathscr{G}'$ est strictement décroissante, on a $j_0 = 0$. Mais alors, une récurrence immédiate montre que $j_i = i$, pour

3. En convenant qu'une suite réduite à un élément est strictement décroissante.

tout $i \in [\![0, r]\!]$. Or, on a

$$G'_s = G = G_r = G'_{j_r} = G'_r,$$

d'où $r = s$. Finalement, $\mathscr{G}' = \mathscr{G}$, et $\mathscr{G}$ est un élément maximal de l'ensemble des suites de composition strictement décroissantes de G. Cela achève la démonstration. □

Ce lemme motive la définition suivante.

3.6. Définition. Soit G un groupe. Une suite de composition de G est une *suite de Jordan-Hölder* si tous ses quotients sont des groupes simples (cf. définition VI-2.18). Autrement dit, une suite de Jordan-Hölder de G est une suite de composition strictement décroissante de G qui ne possède pas de raffinement strict.

Ainsi, si $G = \{1_G\}$, la suite $(\{1_G\})$ est l'unique suite de Jordan-Hölder de G.

3.7. Exemples

(1) Si G est un groupe simple, la suite $(G, \{1_G\})$ est l'unique suite de Jordan-Hölder de G.

(2) Les suites

$$(\mathbb{Z}/6\mathbb{Z}, \langle\overline{3}\rangle, \{\overline{0}\}) \quad \text{et} \quad (\mathbb{Z}/6\mathbb{Z}, \langle\overline{2}\rangle, \{\overline{0}\})$$

sont deux suites de Jordan-Hölder de $\mathbb{Z}/6\mathbb{Z}$.

Les quotients successifs de la première suite sont $\mathbb{Z}/3\mathbb{Z}$ et $\mathbb{Z}/2\mathbb{Z}$, tandis que ceux de la seconde suite sont $\mathbb{Z}/2\mathbb{Z}$ et $\mathbb{Z}/3\mathbb{Z}$.

(3) La suite $(\mathbb{Z}/4\mathbb{Z}, \langle\overline{2}\rangle, \{\overline{0}\})$ est une suite de Jordan-Hölder de $\mathbb{Z}/4\mathbb{Z}$, de quotients successifs $\mathbb{Z}/2\mathbb{Z}$ et $\mathbb{Z}/2\mathbb{Z}$.

(4) La suite $(\mathbb{Z}/2\mathbb{Z}\times\mathbb{Z}/2\mathbb{Z}, \langle(\overline{0}, \overline{1})\rangle, \{(\overline{0}, \overline{0})\})$ est une suite de Jordan-Hölder de $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, de quotients successifs $\mathbb{Z}/2\mathbb{Z}$ et $\mathbb{Z}/2\mathbb{Z}$.

(5) Le groupe $\mathbb{Z}$ n'admet pas de suite de Jordan-Hölder.

En effet, si $(G_0, G_1, \ldots, G_r)$ est une suite de Jordan-Hölder de $\mathbb{Z}$, alors le groupe quotient $G_{r-1}/G_r = G_{r-1}/\{0\} \simeq G_{r-1}$ est simple. Mais, on a $G_{r-1} = n\mathbb{Z}$ pour un entier $n \geqslant 1$ (on ne peut avoir $n = 0$, car un groupe simple n'est pas réduit au neutre). Or, $n\mathbb{Z}$ n'est pas simple, puisque $2n\mathbb{Z}$ en est un sous-groupe non trivial, nécessairement distingué, car $n\mathbb{Z}$ est abélien.

(6) Plus généralement, un groupe abélien infini n'admet pas de suite de Jordan-Hölder.

En effet, les quotients successifs d'une suite de Jordan-Hölder d'un groupe abélien sont des groupes abéliens simples, donc en particulier

finis d'après l'exercice VI-21. En particulier, un groupe abélien possédant une suite de Jordan-Hölder est nécessairement fini.

Les deux derniers exemples montrent qu'un groupe G n'admet pas nécessairement de suite de Jordan-Hölder. Nous n'étudierons pas ici les conditions nécessaires et suffisantes d'existence d'une suite de Jordan-Hölder. En revanche, nous allons donner quelques résultats à ce propos.

3.8. Proposition. *Soit G un groupe non réduit au neutre. Alors :*

(1) *si G admet une suite de Jordan-Hölder, il en est de même pour tout sous-groupe H distingué dans G ;*

(2) *si G admet une suite de Jordan-Hölder, il en est de même pour tout quotient G/H, où H est distingué dans G ;*

(3) *si H est un sous-groupe distingué de G tel que H et G/H admettent une suite de Jordan-Hölder, alors G admet une suite de Jordan-Hölder ;*

(4) *si G est fini, G admet une suite de Jordan-Hölder.*

Démonstration

(1) Soit $(G_0, \ldots, G_r)$ une suite de Jordan-Hölder de G.

Par le lemme 3.3 (1), $(G_0 \cap H, \ldots, G_r \cap H)$ est une suite de composition de H, et l'on a vu que le morphisme de groupes

$$\begin{array}{ccc} G_i \cap H & \longrightarrow & G_i/G_{i+1} \\ g_i & \longmapsto & \overline{g}_i \end{array}$$

induisait par passage au quotient un morphisme de groupes injectif

$$\begin{array}{ccc} \varphi_i : (G_i \cap H)/(G_{i+1} \cap H) & \longrightarrow & G_i/G_{i+1} \\ \tilde{g}_i & \longmapsto & \overline{g}_i. \end{array}$$

L'image de ce morphisme est $G' = \{\overline{g} \mid g \in G_i \cap H\}$, et l'on a un isomorphisme

$$(G_i \cap H)/(G_{i+1} \cap H) \simeq G'.$$

Puisque H est distingué dans G, on voit facilement que G' est distingué dans G_i/G_{i+1}. Comme le groupe G_i/G_{i+1} est simple, on a $G' = \{\overline{1}_G\}$ ou $G' = G_i/G_{i+1}$.

Dans le premier cas, on a $G_i \cap H = G_{i+1} \cap H$, et dans le second le groupe $(G_i \cap H)/(G_{i+1} \cap H)$ est simple. Ainsi, en enlevant les termes qui se répètent dans la suite $(G_0 \cap H, \ldots, G_r \cap H)$, on obtient une suite de Jordan-Hölder de H.

(2) Soit $(G_0, \ldots, G_r)$ une suite de Jordan-Hölder de G. La suite

$$(\langle G_0, H\rangle/H, \ \ldots, \ \langle G_r, H\rangle/H)$$

est alors une suite de composition de G/H, par le lemme 3.3 (2). De plus, $(\langle G_i, H\rangle/H)/(\langle G_{i+1}, H\rangle/H)$ est isomorphe à un quotient du groupe G_i/G_{i+1}. Puisque G_i/G_{i+1} est simple, ses seuls sous-groupes distingués sont $\{\overline{1}_{G_i}\}$ et G_i/G_{i+1}, et les quotients correspondants sont donc respectivement isomorphes à G_i/G_{i+1} et $\{\overline{1}_{G_i}\}$. Dans le second cas, cela signifie que $\langle G_i, H\rangle/H = \langle G_{i+1}, H\rangle/H$. On obtient alors une suite de Jordan-Hölder de G/H en appliquant le lemme 3.3 (2), et en enlevant les termes qui se répètent.

(3) Il suffit d'appliquer le lemme 3.3 (3).

(4) On procède par récurrence sur l'ordre de G. Soit $n = |G|$, et soit (H_n) la proposition suivante :

(H_n) tout groupe d'ordre $\leqslant n$ admet une suite de Jordan-Hölder.

Clairement, (H_1) est vraie. Supposons maintenant que (H_n) soit vraie pour un certain $n \geqslant 2$, et montrons (H_{n+1}).

Si $|G| \leqslant n$, on applique l'hypothèse de récurrence. On peut donc supposer que $|G| = n + 1$. Si G est simple, on a fini puisque $(G, \{1_G\})$ est une suite de Jordan-Hölder de G. Si G n'est pas simple, choisissons un sous-groupe propre H distingué dans G. Alors, G/H est non réduit au neutre, car $H \neq G$, et $|G/H| \leqslant n$, car $H \neq \{1_G\}$. De plus, H est non réduit au neutre et $|H| \leqslant n$, car $H \neq G$. Par hypothèse de récurrence, H et G/H admettent chacun une suite de Jordan-Hölder. Le point (3) montre alors qu'il en est de même pour G. Ainsi, (H_{n+1}) est vraie, et cela achève la récurrence. □

3.9. Remarque. Le point (1) devient faux si H n'est pas supposé distingué dans G (cf. exercice XI-23).

Une suite de Jordan-Hölder n'est pas nécessairement unique, comme le montre déjà l'exemple 3.7 (2). En revanche, nous allons montrer que la suite des quotients successifs l'est, à permutation près des facteurs. La définition suivante sera bien commode.

3.10. Définition. Soit G un groupe.
Deux suites de composition $\mathscr{G} = (G_0, \cdots, G_r)$ et $\mathscr{G}' = (G'_0, \ldots, G'_s)$ sont *équivalentes* si $r = s$ et s'il existe une permutation $\sigma \in \mathfrak{S}(\llbracket 0, r-1 \rrbracket)$ telle que

$$G'_i/G'_{i+1} \simeq G_{\sigma(i)}/G_{\sigma(i)+1}, \qquad \text{pour tout } i \in \llbracket 0, r-1 \rrbracket .$$

3.11. Exemple. Les suites

$$(\mathbb{Z}/6\mathbb{Z}, \langle\overline{3}\rangle, \{\overline{0}\}) \quad \text{et} \quad (\mathbb{Z}/6\mathbb{Z}, \langle\overline{2}\rangle, \{\overline{0}\})$$

sont deux suites de Jordan-Hölder de $\mathbb{Z}/6\mathbb{Z}$ qui sont équivalentes.

Le but de ce qui suit est de démontrer que deux suites de Jordan-Hölder d'un groupe G sont équivalentes (dans le cas où il en existe au moins une). Pour cela, nous allons étudier de plus près la relation d'ordre induite par le raffinement.

On commence par un résultat dû à Zassenhaus.

3.12. Théorème (Lemme du papillon)
Soit G un groupe, et soient quatre sous-groupes H, H', K et K' de G tels que $H' \lhd H$ et $K' \lhd K$. Alors :

(1) *les ensembles $H'(H \cap K'), H'(H \cap K), (H' \cap K)K'$ et $(H \cap K)K'$ sont des sous-groupes de G ;*

(2) *l'ensemble $H'(H \cap K')$ est un sous-groupe distingué de $H'(H \cap K)$;*

(3) *l'ensemble $(H' \cap K)K'$ est un sous-groupe distingué de $(H \cap K)K'$;*

(4) *on a un isomorphisme de groupes*

$$H'(H \cap K)/H'(H \cap K') \simeq (H \cap K)K'/(H' \cap K)K'.$$

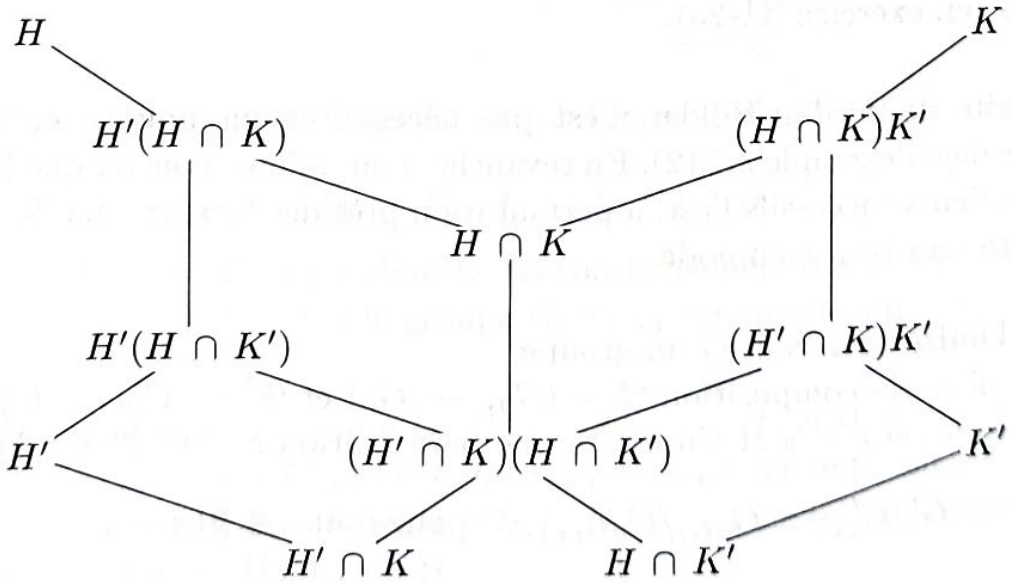
3.13. Remarque

On peut rassembler les nombreux sous-groupes intervenant dans la démonstration suivante dans un diagramme, les traits représentant l'inclusion.

En tenant compte de l'égalité

$$(H \cap K')(H' \cap K) = (H' \cap K)(H \cap K'),$$

on obtient le dessin suivant, qui explique le nom de « lemme du papillon ».



Démonstration

(1) Notons que H', $H \cap K'$ et $H \cap K$ sont des sous-groupes de G. Comme H' est distingué dans H, $H'(H \cap K')$ et $H'(H \cap K)$ sont des sous-groupes de H par le théorème 2.5. Ce même théorème nous dit également que les ensembles $K'(H' \cap K)$ et $K'(H \cap K)$ sont des sous-groupes de G, et que ces sous-groupes sont aussi égaux à $(H' \cap K)K'$ et $(H \cap K)K'$.

(2) Comme $H'(H \cap K')$ est un sous-groupe de G, inclus dans le sous-groupe $H'(H \cap K)$, c'est aussi un sous-groupe de $H'(H \cap K)$. Comme K' est distingué dans K, on montre alors aisément que $H \cap K'$ est distingué dans $H \cap K$.

De plus, H' étant distingué dans H, il est en particulier stable par conjugaison d'un élément de $H \cap K$.

Soient maintenant $h' \in H'$ et $k \in H \cap K$. On a alors

$$(h'k)H'(h'k)^{-1} = h'(kH'k^{-1})(h')^{-1} \subset h'H'(h')^{-1} \subset H' \subset H'(H \cap K').$$

De plus, on a également

$$(h'k)(H \cap K')(h'k)^{-1} = h'\big(k(H \cap K')k^{-1}\big)(h')^{-1} \subset h'(H \cap K')(h')^{-1}.$$

Or, $H \cap K'$ étant distingué dans $H \cap K$, on a $H'(H \cap K') = (H \cap K')H'$ par le théorème 2.5, appliqué aux sous-groupes $H \cap K'$ et H' de $H \cap K$. On en déduit aisément l'inclusion

$$(h'k)(H \cap K')(h'k)^{-1} \subset H \cap K'.$$

Mais on a

$$(h'k)H'(H \cap K')(h'k)^{-1} = \big((h'k)H'(h'k)^{-1}\big)\big((h'k)(H \cap K')(h'k)^{-1}\big),$$

et par conséquent

$$(h'k)H'(H \cap K')(h'k)^{-1} \subset \big(H'(H \cap K')\big)\big(H'(H \cap K')\big) \subset H'(H \cap K'),$$

d'où le résultat.

(3) En échangeant les rôles de H et K, et de H' et K', le point (2) et le théorème 2.5 montrent que $(H' \cap K)K' = K'(H' \cap K)$ est un sous-groupe distingué de $(H \cap K)K' = K'(H \cap K)$.

(4) Nous appliquons le théorème 2.5 aux deux sous-groupes $H'(H \cap K')$ et $H \cap K$ de $H'(H \cap K)$, ce que l'on peut faire grâce au point (2).

Remarquons tout d'abord que l'on a

$$\big(H'(H \cap K')\big)(H \cap K) = H'(H \cap K).$$

En effet, on a d'une part

$$\big(H'(H \cap K')\big)(H \cap K) = H'\big((H \cap K')(H \cap K)\big) \subset H'(H \cap K),$$

et d'autre part l'inclusion $H' \subset H'(H \cap K)$ implique que l'on a

$$H'(H \cap K) \subset \big(H'(H \cap K')\big)(H \cap K).$$

Montrons maintenant l'égalité

$$\big(H'(H \cap K')\big) \cap (H \cap K) = (H' \cap K)(H \cap K').$$

Clairement, $H' \cap K$ est inclus dans $H \cap K$. Il est aussi inclus dans H', donc dans $H'(H \cap K')$. De même, $H \cap K'$ est inclus dans l'intersection. Comme $\big(H'(H \cap K')\big) \cap (H \cap K)$ est stable par produit, on en déduit que

$$(H' \cap K)(H \cap K') \subset \big(H'(H \cap K')\big) \cap (H \cap K).$$

Soit maintenant $x \in \big(H'(H \cap K')\big) \cap (H \cap K)$. On a donc $x \in H \cap K$, ainsi que $x = h'k'$, avec $h' \in H'$ et $k' \in H \cap K'$. On a alors $h' = x(k')^{-1}$. Comme on a $x \in K$ et $k' \in K' \subset K$, on en déduit que $h' \in H' \cap K$. Ainsi, on a $x \in (H' \cap K)(H \cap K')$, d'où l'égalité annoncée.

Le théorème 2.5 fournit alors un isomorphisme de groupes

$$H'(H \cap K)/H'(H \cap K') \simeq (H \cap K)/\big((H' \cap K)(H \cap K')\big).$$

En échangeant les rôles de H et K, et de H' et K', et en utilisant le théorème 2.5, on obtient aussi l'isomorphisme de groupes

$$(K \cap H)K'/(K \cap H')K' \simeq (K \cap H)/\big((K' \cap H)(K \cap H')\big),$$

Le théorème 2.5, appliqué cette fois aux sous-groupes $H' \cap K$ et $H \cap K'$ de $H \cap K$, montre que $(H \cap K')(H' \cap K) = (H' \cap K)(H \cap K')$, puisque le sous-groupe $H \cap K'$ est distingué dans $H \cap K$, par exemple.

On obtient alors finalement

$$(H \cap K)K'/(H' \cap K)K' \simeq (H \cap K)/\big((H' \cap K)(H \cap K')\big).$$

Ce dernier groupe étant isomorphe au quotient $H'(H \cap K)/H'(H \cap K')$, on a le résultat voulu. Cela achève la démonstration. □

Nous allons pouvoir maintenant énoncer et démontrer le théorème du raffinement de Schreier.

3.14. Théorème (Schreier). *Soit G un groupe. Pour toutes suites de composition $\mathscr{H} = (H_0, \ldots, H_r)$ et $\mathscr{K} = (K_0, \ldots, K_s)$, il existe des raffinements $\mathscr{H}'$ et $\mathscr{K}'$ de $\mathscr{H}$ et $\mathscr{K}$ respectivement qui sont équivalents.*

Démonstration. Pour tout $i \in [\![0, r-1]\!]$, et pour tout $j \in [\![0, s]\!]$, on pose

$$H_{i,j} = H_{i+1}(H_i \cap K_j).$$

Puisque K_{j+1} est un sous-groupe de K_j, $H_{i,j+1}$ est un sous-groupe de $H_{i,j}$.

D'après le point (2) du lemme du papillon, $H_{i,j+1} = H_{i+1}(H_i \cap K_{j+1})$ est un sous-groupe distingué de $H_{i,j} = H_{i+1}(H_i \cap K_j)$ pour tout $j \in [\![0, s-1]\!]$. De plus, on a $H_{i,0} = H_{i+1}H_i = H_i$, ainsi que $H_{i,s} = H_{i+1}$, car $K_0 = G$ et $K_s = \{1_G\}$. On obtient alors une suite de sous-groupes $\mathscr{H}'$ de G en ordonnant les sous-groupes $H_{i,j}$, avec $i \in [\![0, r-1]\!]$ et $j \in [\![0, s-1]\!]$, grâce à l'ordre lexicographique, suivi du groupe trivial. Autrement dit, on a

$$\mathscr{H}' = (H_{0,0}, \ldots, H_{0,s-1}, H_{1,0}, \ldots, H_{1,s-1}, \ldots, H_{r-1,0}, \ldots, H_{r-1,s-1}, \{1_G\}).$$

Notons que c'est bien une suite de composition de G, puisque pour tout entier $i \in [\![0, r-1]\!]$, le sous-groupe $H_{i+1,0} = H_{i+1} = H_{i,s}$ est distingué dans $H_i = H_{i,s-1}$, $\{1_G\}$ est distingué dans $H_{r-1,s-1}$, et $H_{0,0} = H_0 = G$. De plus, on a $rs+1$ sous-groupes, si bien que $\mathscr{H}'$ est de longueur rs. Enfin, la suite $\mathscr{H}'$ est un raffinement de $\mathscr{H}$ par définition.

Notons que $H_{i,s} = H_{i+1} = H_{i+1,0}$, si bien que le sous-groupe $H_{i,j}$ dans la suite de composition est suivi par $H_{i,j+1}$, et cela même si $j = s-1$. Les quotients de $\mathscr{H}'$ sont donc les groupes

$$H_{i,j}/H_{i,j+1} = H_{i+1}(H_i \cap K_j)/H_{i+1}(H_i \cap K_{j+1}),$$

où $(i,j) \in [\![0, r-1]\!] \times [\![0, s-1]\!]$.

Pour tout $i \in [\![0, r]\!]$, et pour tout $j \in [\![0, s-1]\!]$, on pose maintenant

$$K_{j,i} = (H_i \cap K_j)K_{j+1}.$$

On vérifie alors comme précédemment que la suite de sous-groupes

$$\mathscr{K}' = (K_{0,0}, \ldots, K_{0,r-1}, K_{1,0}, \ldots, K_{1,r-1}, \ldots, K_{s-1,0}, \ldots, K_{s-1,r-1}, \{1_G\})$$

est un raffinement de $\mathscr{K}$, de longueur rs, et que le sous-groupe $K_{j,i}$ dans la suite de composition est suivi par $K_{j,i+1}$, et cela même si $i = r-1$. Les quotients de $\mathscr{K}'$ sont donc les groupes

$$K_{j,i}/K_{j,i+1} = (H_i \cap K_j)K_{j+1}/(H_{i+1} \cap K_j)K_{j+1},$$

où $(i,j) \in [\![0, r-1]\!] \times [\![0, s-1]\!]$.

D'après le lemme du papillon, on a

$$K_{j,i}/K_{j,i+1} \simeq H_{i,j}/H_{i,j+1}, \quad \text{pour tout } (i,j) \in [\![0, r-1]\!] \times [\![0, s-1]\!].$$

Les deux raffinements sont donc équivalents, ce qui achève la démonstration. □

3.15. Remarque. Dans la démonstration, les raffinements proposés ne sont pas indexés par l'intervalle $[\![0, rs]\!]$, et l'équivalence n'est pas donnée par une permutation des éléments de cet ensemble. Une renumérotation idoine des sous-groupes permet de lever cette imprécision.

En effet, on peut vérifier aisément que les entiers

$$is + j, \quad (i,j) \in [\![0, r-1]\!] \times [\![0, s-1]\!]$$

sont deux à deux distincts, et décrivent l'ensemble $[\![0, rs-1]\!]$. De même, les entiers

$$jr + i, \quad (i,j) \in [\![0, r-1]\!] \times [\![0, s-1]\!]$$

sont deux à deux distincts, et décrivent l'ensemble $[\![0, rs-1]\!]$.

Pour tout $(i,j) \in [\![0, r-1]\!] \times [\![0, s-1]\!]$, on pose alors

$$H'_{is+j} = H_{i,j} \quad \text{et} \quad K'_{jr+i} = K_{j,i}.$$

On pose également $H'_{rs} = K'_{rs} = \{1_G\}$. Notons σ l'unique permutation de l'ensemble $[\![0, rs]\!]$ fixant rs et telle que

$$\sigma(is + j) = jr + i, \quad \text{pour tout } (i,j) \in [\![0, r-1]\!] \times [\![0, s-1]\!].$$

Alors, les suites $(H'_0, \ldots, H'_{rs})$ et $(K'_0, \ldots, K'_{rs})$ sont des raffinements de $\mathscr{H}$ et $\mathscr{K}$ respectivement, et vérifiant

$$K'_m / K'_{m+1} \simeq H'_{\sigma(m)} / H'_{\sigma(m)+1}, \quad \text{pour tout } m \in [\![0, rs-1]\!].$$

Les détails sont laissés au lecteur.

Nous pouvons maintenant démontrer le théorème suivant.

3.16. Théorème (Jordan-Hölder). *Soit G un groupe admettant au moins une suite de Jordan-Hölder. Alors, deux suites de Jordan-Hölder de G sont équivalentes.*

Démonstration. Soient $\mathscr{H}$ et $\mathscr{K}$ deux suites de Jordan-Hölder de G. Par le théorème de Schreier, il existe des raffinements $\mathscr{H}'$ et $\mathscr{K}'$ de $\mathscr{H}$ et $\mathscr{K}$ respectivement qui sont équivalents. D'après le lemme 3.5, une suite de Jordan-Hölder n'admet pas de raffinement strict. On a donc $\mathscr{H}' = \mathscr{H}$, ainsi que $\mathscr{K}' = \mathscr{K}$. Mais alors, $\mathscr{H}$ et $\mathscr{K}$ sont équivalentes, d'où le résultat. □

3.17. Remarques

(1) Si G et G' sont deux groupes isomorphes, alors G admet une suite de Jordan-Hölder si, et seulement si, il en est de même de G'. Dans ce cas, les listes des quotients successifs de leurs suites de Jordan-Hölder sont les mêmes, à isomorphisme et permutation près des éléments de ces listes.

(2) Dans le cas des groupes finis, on peut considérer le théorème de Jordan-Hölder comme un analogue du théorème fondamental de l'arithmétique : à tout groupe fini G, on associe une suite de groupes finis simples, et une telle suite est unique à permutation près, et à isomorphisme près des éléments de la suite.

Cette analogie peut être de plus légitimée par le fait que, dans le cas abélien, les groupes simples sont exactement les $\mathbb{Z}/p\mathbb{Z}$, avec p premier. Malheureusement, cette analogie s'arrête là. Si l'on peut reconstituer un entier positif à partir de la liste de ses diviseurs premiers (avec répétition), il n'en est pas de même pour les groupes finis, même dans le cas abélien, comme le montre les exemples 3.7 (3) et (4). Même si la classification complète des groupes finis simples est maintenant connue, elle ne suffit pas en elle-même pour pouvoir classifier les groupes finis à partir d'une suite de Jordan-Hölder. En fait, il faut aussi savoir résoudre le problème des extensions de groupes : étant donné deux groupes N et K, trouver tous les groupes G (à isomorphisme près) ayant un sous-groupe H distingué dans G, isomorphe à N, et tel que l'on ait un isomorphisme $G/H \simeq K$.

Nous allons maintenant nous intéresser à un autre type de suites de composition, à savoir celles dont les quotients successifs sont abéliens ou cycliques. Cela conduit à la notion de groupe résoluble, notion historiquement très importante puisqu'elle est liée au problème de résolution d'équations polynomiales par radicaux et aux travaux d'Abel et Galois.

3.18. Définition. On dit qu'un groupe G est *résoluble* s'il admet une suite de composition $(G_0, \ldots, G_r)$ dont les quotients successifs sont abéliens.

On dit qu'un groupe G est *super-résoluble* s'il admet une suite de composition $(G_0, \ldots, G_r)$ dont les quotients successifs sont cycliques, et où les sous-groupes $G_0, \ldots, G_r$ sont distingués dans G.

3.19. Remarque. On peut démontrer qu'un groupe résoluble admet une suite de composition $(G_0, \ldots, G_r)$ dont les quotients successifs sont abéliens, et où $G_0, \ldots, G_r$ sont distingués dans G (cf. exercice 23). On peut également démontrer qu'un groupe résoluble admet une suite de composition $(G_0, \ldots, G_r)$ dont les quotients successifs sont cycliques, mais dans ce cas les sous-groupes $G_0, \ldots, G_r$ ne sont pas nécessairement distingués dans le groupe G (cf. exercice XII-5).

3.20. Exemples

(1) Tout groupe super-résoluble est bien entendu résoluble.

(2) Tout groupe simple non abélien n'est pas résoluble.

(3) Tout groupe abélien fini G est super-résoluble. En effet, tout sous-groupe de G est alors distingué dans G, et les quotients successifs d'une suite de Jordan-Hölder sont des groupes abéliens finis simples, donc isomorphes à $\mathbb{Z}/p\mathbb{Z}$, avec p premier d'après l'exercice VI-21.

(4) Le groupe $\mathfrak{A}_4$ est résoluble, mais pas super-résoluble.

En effet, le sous-groupe $H = \{\mathrm{Id},\ (1\ 2)(3\ 4),\ (1\ 3)(2\ 4),\ (1\ 4)(2\ 3)\}$ est distingué dans $\mathfrak{A}_4$, isomorphe à $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$. Mais alors, $\mathfrak{A}_4/H$ est d'ordre 3, et est donc isomorphe à $\mathbb{Z}/3\mathbb{Z}$. Ainsi, la suite $(\mathfrak{A}_4,\ H,\ \{\mathrm{Id}\})$ est une suite de composition dont les quotients successifs sont abéliens, et $\mathfrak{A}_4$ est résoluble.

En revanche, $\mathfrak{A}_4$ n'est pas super-résoluble. En effet, les seuls sous-groupes distingués de $\mathfrak{A}_4$ sont $\{\mathrm{Id}\}, H$ et $\mathfrak{A}_4$. Toute suite de composition de $\mathfrak{A}_4$ dont les sous-groupes sont distingués dans $\mathfrak{A}_4$ est donc de la forme

$$(\mathfrak{A}_4,\ \ldots,\ \mathfrak{A}_4,\ H,\ \ldots,\ H,\ \{\mathrm{Id}\},\ \ldots,\ \{\mathrm{Id}\}).$$

Aucune d'entre elles n'ayant des quotients successifs tous cycliques, $\mathfrak{A}_4$ n'est pas super-résoluble.

(5) Pour tout $n \geqslant 5$, $\mathfrak{S}_n$ n'est pas résoluble, puisque les seules suites de composition possibles de $\mathfrak{S}_n$ ne comportant pas de répétitions sont les suites $(\mathfrak{S}_n, \{\mathrm{Id}\})$ et $(\mathfrak{S}_n, \mathfrak{A}_n, \{\mathrm{Id}\})$, d'après le corollaire VIII-5.7.

(6) Citons le résultat spectaculaire de Feit et Thompson, dont la démonstration fait plus de trois cents pages : tout groupe fini d'ordre impair est résoluble. En particulier, un groupe non abélien d'ordre impair n'est pas simple.

3.21. Proposition. *Soit G un groupe. Alors :*

(1) *si G est (super-)résoluble, il en est de même de tout sous-groupe de G ;*

(2) *si G est (super-)résoluble, il en est de même de tout quotient de G ;*

(3) *pour tout sous-groupe H distingué dans G, le groupe G est résoluble si, et seulement si, les groupes H et G/H sont résolubles.*

Si de plus H est un sous-groupe de $Z(G)$ et si G/H est super-résoluble, alors G est super-résoluble.

Démonstration

(1) Soit $(G_0, \ldots, G_r)$ une suite de composition de G et soit H un sous-groupe de G.

Alors, par le lemme 3.3 (1), $(G_0 \cap H, \ldots, G_r \cap H)$ est une suite de composition de H. De plus, pour tout $i \in [\![0, r-1]\!]$, $G_i \cap H/G_{i+1} \cap H$ est isomorphe à un sous-groupe de G_i/G_{i+1}. Si G_i/G_{i+1} est abélien (resp. cyclique), il en est de même de tout sous-groupe, et donc du groupe $G_i \cap H/G_{i+1} \cap H$(cf. théorème VI-6.5 pour le cas cyclique). De plus, si G_i est distingué dans G, alors $G_i \cap H$ est distingué dans H. Cela montre que si G est (super-)résoluble, il en est de même de H.

(2) Soit $(G_0, \ldots, G_r)$ une suite de composition de G et soit H un sous-groupe distingué de G.

Alors, par le lemme 3.3 (2), $(\langle G_0, H\rangle/H, \ldots, \langle G_r, H\rangle/H)$ est une suite de composition de G/H. De plus, pour tout $i \in [\![0, r-1]\!]$, le groupe quotient $(\langle G_i, H\rangle/H)/(\langle G_{i+1}, H\rangle/H)$ est isomorphe à un sous-groupe de G_i/G_{i+1}.

Si G_i/G_{i+1} est abélien (resp. cyclique), il en est de même de tout quotient, et donc de $(\langle G_i, H\rangle/H)/(\langle G_{i+1}, H\rangle/H)$. Si de plus G_i est distingué dans G, comme H est aussi distingué dans G, $\langle G_i, H\rangle$ est distingué dans G, et $\langle G_i, H\rangle/H$ est alors distingué dans G/H, d'où le résultat.

(3) Il reste à démontrer le sens indirect. Pour cela, il suffit d'appliquer le lemme 3.3 (3), et de remarquer que si de plus H est un sous-groupe de $Z(G)$, tous les sous-groupes d'une suite de composition de H sont distingués dans G. □

3.22. Remarque. La première assertion du point (3) est fausse en toute généralité pour les groupes super-résolubles. En effet, le sous-groupe H de $\mathfrak{A}_4$ de l'exemple 3.20 (4) est abélien, donc super-résoluble[4] d'après l'exemple 3.20 (3). De plus, $\mathfrak{A}_4/H$ est isomorphe à $\mathbb{Z}/3\mathbb{Z}$, qui est aussi super-résoluble. En revanche, toujours d'après l'exemple 3.20 (4), $\mathfrak{A}_4$ n'est pas super-résoluble.

Nous allons maintenant nous intéresser aux p-groupes.

3.23. Théorème. *Soit G un groupe d'ordre p^n, avec p premier, et $n \geqslant 0$. Alors, il existe au moins une suite de composition $(G_0, \ldots, G_n)$, où G_i est un sous-groupe distingué de G d'indice p^i, pour tout $i \in [\![0, n]\!]$.*
En particulier, tout p-groupe fini G est super-résoluble (donc résoluble).

Démonstration. Soit G un groupe d'ordre p^n, avec p premier, et $n \geqslant 0$. Le cas $n = 0$ étant évident, on peut supposer que $n \geqslant 1$.
On procède par récurrence sur n. Soit (H_n) la propriété :

(H_n) Le théorème est vrai pour tout groupe d'ordre p^n.

Si $n = 1$, le groupe G est cyclique d'ordre p, et la propriété (H_1) est vraie. Supposons maintenant que (H_n) soit vraie pour un entier $n \geqslant 1$, et montrons (H_{n+1}). Soit G un groupe d'ordre p^{n+1}. Son centre $Z(G)$ est donc non trivial d'après la proposition VII-2.7. En particulier, son ordre est divisible par p, et le théorème de Cauchy (théorème VII-2.3) assure l'existence d'un élément $x \in Z(G)$ d'ordre p.

4. On peut le voir directement ici : le sous-groupe $K = \langle (1\ 2)\rangle$ est cyclique d'ordre 2, et $H/K \simeq \mathbb{Z}/2\mathbb{Z}$.

Alors, $\langle x \rangle$ est cyclique d'ordre p, et distingué dans G, car $x \in Z(G)$. Mais, alors, le groupe $G/\langle x \rangle$ est d'ordre p^n. Par hypothèse de récurrence, il existe une suite de composition $(G_0/\langle x \rangle, \ldots, G_n/\langle x \rangle)$ de $G/\langle x \rangle$ telle que $G_i/\langle x \rangle$ soit un sous-groupe distingué de $G/\langle x \rangle$ d'indice p^i (où G_i est un sous-groupe de G contenant $\langle x \rangle$), et cela pour tout $i \in [\![0, n]\!]$. Puisque le sous-groupe $G_i/\langle x \rangle$ est distingué dans $G/\langle x \rangle$, G_i est distingué dans G par la proposition 1.6. De plus, on a les égalités

$$[G : G_i] = \frac{|G|}{|G_i|} = \frac{|G|}{|\langle x \rangle|} \times \frac{|\langle x \rangle|}{|G_i|} = [G/\langle x \rangle : G_i/\langle x \rangle] = p^i.$$

Puisque $G_n/\langle x \rangle$ est trivial, on a $G_n = \langle x \rangle$, qui est cyclique d'ordre p. Si l'on pose $G_{n+1} = \{1_G\}$, la suite de composition $(G_0, \ldots, G_{n+1})$ est une suite de composition de G qui vérifie les conditions demandées. Cela achève la récurrence.

Le dernier point découle du fait que, pour tout $i \in [\![0, n]\!]$, G_i/G_{i+1} est d'ordre p premier, donc cyclique. □

3.24. Remarque. Ce théorème montre qu'une suite de Jordan-Hölder d'un groupe d'ordre p^n a tous ses quotients successifs isomorphes à $\mathbb{Z}/p\mathbb{Z}$. Ainsi, les suites de Jordan-Hölder ne permettent pas de distinguer les groupes d'ordre p^n.

4. Présentations de groupes

Nous allons maintenant expliquer comment les quotients permettent de construire des groupes vérifiant des relations données, et aussi comment déterminer (au moins en théorie) tous les morphismes d'un groupe G donné dans un groupe G'.

Avant de rentrer dans le vif du sujet, commençons par un exemple. Considérons le groupe D_4 des isométries vectorielles préservant un carré de centre O. Ce groupe est engendré par un quart de tour σ, et une réflexion orthogonale τ (par rapport à une diagonale, par exemple). Ces deux éléments vérifient

$$\sigma^4 = \mathrm{Id}, \quad \tau^2 = \mathrm{Id}, \quad \sigma\tau\sigma^{-1} = \sigma^{-1}.$$

Remarquons qu'il existe d'autres groupes possédant des éléments vérifiant ces relations. Par exemple, si $G = \{\pm 1\} \times \{\pm 1\}$, les éléments $a = (1, -1)$ et $b = (-1, 1)$ engendrent G, et vérifient également

$$a^4 = 1_G, \quad b^2 = 1_G, \quad aba^{-1} = a^{-1}.$$

En revanche, D_4 et ses deux générateurs σ et τ vérifient la propriété suivante, contrairement à $\{\pm 1\} \times \{\pm 1\}$: pour tout groupe G' possédant deux

éléments r, $s \in G'$ vérifiant

$$r^4 = 1_{G'}, \quad s^2 = 1_{G'}, \quad srs^{-1} = r^{-1},$$

il existe un unique morphisme de groupes $f : D_4 \longrightarrow G'$ tel que $f(\sigma) = r$ et $f(\tau) = s$.

Pour s'en convaincre, il suffit de remarquer que tout élément de D_4 s'écrit de manière unique sous la forme

$$\sigma^i \tau^j, \quad i \in [\![0, 3]\!], \ j = 0, 1.$$

Il suffit alors de poser

$$f(\sigma^i \tau^j) = r^i s^j, \quad \text{pour tous } i \in [\![0, 3]\!], \ j = 0, 1,$$

et de vérifier que l'on obtient bien un morphisme de groupes, ce qui est un calcul laissé au lecteur. Remarquons que, inversement, si $f : D_4 \longrightarrow G'$ est un morphisme de groupes, alors on a

$$f(\sigma)^4 = f(\sigma^4) = f(\mathrm{Id}) = 1_{G'},$$

et de même

$$f(\tau)^2 = \mathrm{Id}, \quad \text{et} \quad f(\tau) f(\sigma) f(\tau)^{-1} = f(\sigma)^{-1}.$$

Ainsi, l'ensemble $\mathrm{Hom}(D_4, G')$ est en bijection avec les éléments r, $s \in G'$ vérifiant

$$r^4 = 1, \quad s^2 = 1, \quad srs^{-1} = r^{-1}.$$

Autrement dit, pour qu'il existe au moins un morphisme $D_4 \longrightarrow G'$, il faut et il suffit que les relations vérifiées dans D_4 soient aussi vérifiées dans G'.

Notons que $\{\pm 1\} \times \{\pm 1\}$ et ses deux générateurs a, b ne jouissent pas de la même propriété.

Par exemple, il n'existe pas de morphisme de groupes

$$f : \{\pm 1\} \times \{\pm 1\} \longrightarrow D_4$$

tel que $f(a) = \sigma$ et $f(b) = \tau$. En effet, on a

$$f(a)^2 = f(a^2) = f\big((1, 1)\big) = 1_{D_4},$$

mais $\sigma^2 \neq \mathrm{Id}$.

Remarquons la propriété évoquée ci-dessus et une utilisation du premier théorème d'isomorphisme donnent en particulier le résultat suivant : tout groupe G' engendré par deux éléments r, s vérifiant les relations

$$r^4 = 1_{G'}, \quad s^2 = 1_{G'}, \quad srs^{-1} = r^{-1}$$

est isomorphe à un quotient de D_4. En quelque sorte, D_4 est « le groupe le plus général » vérifiant les relations ci-dessus.

Nous allons maintenant expliquer de manière informelle comment construire un groupe le plus général possible vérifiant des relations imposées à l'avance.

Remarquons que si G est engendré par une famille d'éléments X, tout élément de G peut s'écrire sous la forme

$$x_1^{n_1} \cdots x_k^{n_k}, \quad k \geqslant 0,\ n_i \in \mathbb{Z} \setminus \{0\},\ x_i \in X,$$

avec $x_i \neq x_{i+1}$, pour tout $i \in [\![1, k-1]\!]$.

Bien sûr, cette écriture est loin d'être unique. L'idée est d'utiliser un groupe « universel » F_X, engendré par une certaine partie X, telle que l'écriture précédente soit unique. En particulier, dans un tel groupe, il n'y a aucune relation non triviale entre les éléments de X : pour tout $k \geqslant 0$, pour tous entiers $n_1, \ldots, n_k \in \mathbb{Z}$, et pour tous $x_1, \ldots, x_k \in X$ tels que $x_i \neq x_{i+1}$ pour tout $i \in [\![1, k-1]\!]$, on a

$$x_1^{n_1} \cdots x_k^{n_k} = 1_{F_X} \quad \Longrightarrow \quad n_i = 0, \ \text{ pour tout } i \in [\![1, n]\!].$$

On peut alors démontrer qu'un tel groupe F_X possède la propriété suivante : *pour tout groupe G', et toute famille $(a_x)_{x \in X}$ d'éléments de G', il existe un unique morphisme $f : F_X \longrightarrow G'$, tel que $f(x) = a_x$, pour tout $x \in X$.*

Nous le ferons d'ailleurs plus loin. Lorsque X est une partie à n éléments, on dira que F_X est un groupe libre sur n générateurs.

Une fois l'existence de ce groupe F_X acquise, pour obtenir un groupe universel vérifiant des relations données, il suffira de quotienter par le sous-groupe distingué de F_X « engendré » par les dites relations.

Par exemple, supposons qu'un groupe F_2 libre sur deux générateurs g_1, g_2 existe, et soit H l'intersection des sous-groupes distingués de F_2 contenant les éléments g_1^4, g_2^2 et $g_2 g_1 g_2^{-1} g_1$.

Par suite, on vérifie aisément qu'un morphisme de groupes $\varphi : F_2 \longrightarrow G'$ vérifie $H \subset \mathrm{Ker}(\varphi)$ si, et seulement si, g_1^4, g_2^2 et $g_2 g_1 g_2^{-1} g_1 \in \mathrm{Ker}(\varphi)$. À l'aide du théorème de factorisation, on en déduit alors que F_2/H possède la même propriété universelle que D_4. On montrera d'ailleurs que $F_2/H \simeq D_4$.

Nous allons maintenant formaliser tout ce qui précède.

4.1. Définition. Soit G un groupe, et soit X une partie de G.
On dit que X est *libre*[5] si pour tout $k \geqslant 1$, pour tous $m_1, \ldots, m_k \in \mathbb{Z} \setminus \{0\}$, et pour tous $x_1, \ldots, x_k \in X$ vérifiant $x_j \neq x_{j+1}$ pour tout $j \in [\![1, k-1]\!]$, on a

$$x_1^{m_1} \cdots x_k^{m_k} \neq 1_G.$$

Une partie de G qui n'est pas libre est dite *liée*. On dit que X est une *base* de G si elle est libre et génératrice.

5. De manière informelle, une partie de G est libre s'il n'y a pas de relations non triviales entre ses éléments.

4.2. Lemme. *Soit G un groupe, et soit X une partie de G. Alors, X est une base de G si, et seulement si, tout élément $g \in G$ s'écrit de manière unique sous la forme*

$$g = x_1^{n_1} \cdots x_k^{n_k}, \quad k \geqslant 0,\ n_i \in \mathbb{Z} \setminus \{0\},\ x_i \in X,$$

avec $x_i \neq x_{i+1}$, pour tout $i \in [\![1, k-1]\!]$.

Démonstration. Le sens indirect étant clair, montrons le sens direct. Une base X étant une famille génératrice, il reste à vérifier que si

$$x_1^{n_1} \cdots x_k^{n_k} = y_1^{m_1} \cdots y_\ell^{m_\ell},$$

alors on a $k = \ell, n_i = m_i$ et $x_i = y_i$ pour tout $i \in [\![1, k]\!]$ (avec des notations évidentes).

On procède par récurrence sur l'entier ℓ. Pour $\ell = 0$, le membre de droite est égal à 1_G, et le résultat provient de la liberté de la famille X. Supposons que le résultat soit vrai au rang $\ell \geqslant 0$ et montrons-le au rang $\ell+1$. Considérons une égalité de la forme

$$x_1^{n_1} \cdots x_k^{n_k} = y_1^{m_1} \cdots y_\ell^{m_\ell} y_{\ell+1}^{m_{\ell+1}}.$$

L'égalité précédente se récrit

$$x_1^{n_1} \cdots x_{k-1}^{n_{k-1}} x_k^{n_k} y_{\ell+1}^{-m_{\ell+1}} = y_1^{m_1} \cdots y_\ell^{m_\ell}.$$

Si $y_{\ell+1} \neq x_k$, l'hypothèse de récurrence entraîne alors que $k+1 = \ell$, ainsi que $y_\ell = y_{\ell+1}$, d'où une contradiction.

On a donc $y_{\ell+1} = x_k$, et l'égalité

$$x_1^{n_1} \cdots x_{k-1}^{n_{k-1}} x_k^{n_k - m_{\ell+1}} = y_1^{m_1} \cdots y_\ell^{m_\ell}.$$

Si $n_k - m_{\ell+1} \neq 0$, l'hypothèse de récurrence nous permet d'en déduire en particulier que $k = \ell$, et que $y_k = x_k = y_{k+1}$, ce qui contredit donc le choix de $y_1, \ldots, y_\ell$. Ainsi, $n_k = m_{\ell+1}$, et par conséquent

$$x_1^{n_1} \cdots x_{k-1}^{n_{k-1}} = y_1^{m_1} \cdots y_\ell^{m_\ell}.$$

On applique alors à nouveau l'hypothèse de récurrence pour achever le raisonnement. □

Maintenant que l'on a la notion de base, on peut introduire la notion de groupe libre.

4.3. Définition. On dit que un groupe F est *libre* s'il possède au moins une base. Si X est un ensemble, un *groupe libre sur X* est un couple (F, ι), où F est un groupe, et $\iota : X \hookrightarrow F$ est une application injective telle que l'ensemble $\iota(X)$ soit une base de F.

On dit que deux groupes libres (F, ι) et (F', ι') sur X sont *isomorphes* s'il existe un isomorphisme de groupes $\varphi : F \xrightarrow{\sim} F'$ tel que

$$\varphi\big(\iota(x)\big) = \iota'(x), \quad \text{pour tout } x \in X.$$

4.4. Remarque. Cette définition dit en substance la chose suivante : un groupe libre sur X est un groupe qui contient une base Y équipotente à X. Néanmoins, pour les calculs, on a parfois besoin de spécifier la base, i.e. la bijection de X sur Y, bijection qui est justement donnée par ι. Autrement dit, un groupe libre sur X est la donnée d'un groupe libre et d'une base équipotente à X.

Deux groupes libres sur X sont alors isomorphes s'il existe un isomorphisme de groupes qui envoie la base du premier groupe sur la base du second.

4.5. Exemples

(1) Le groupe trivial est libre, de base vide.

(2) Un groupe fini G non trivial n'est jamais libre.
En effet, une base X de G contiendrait un élément x non trivial. Mais alors, on aurait la relation $x^{o(x)} = 1_G$, et X ne serait pas libre, d'où une contradiction.

(3) Le groupe $\mathbb{Z}$ est libre, puisque $\{1\}$ en est une base. Ainsi, si $X = \{x\}$ est un singleton, et si $\iota : X \hookrightarrow \mathbb{Z}$ est l'application qui envoie x sur 1, le couple $(\mathbb{Z}, \iota)$ est un groupe libre sur X.

(4) Soient $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$; soit G le sous-groupe de $\mathrm{GL}_2(\mathbb{C})$ engendré par A et B. Alors, $\{A, B\}$ est une base de G.
Par définition, $\{A, B\}$ est génératrice. Il faut donc vérifier qu'elle est libre. Considérons une matrice de la forme

$$M = X_1^{m_1} \cdots X_r^{m_r}, \quad r \geqslant 1,\ m_j \in \mathbb{Z} \setminus \{0\},\ X_j \in \{A, B\},$$

avec $X_i \neq X_{i+1}$, pour tout $i \in [\![1, r-1]\!]$. Il faut s'assurer que $M \neq \mathrm{I}_2$. Quitte à remplacer M par AMA^{-1}, on peut supposer que $X_1 = A$.
Pour tout $j \in [\![0, r]\!]$, on note $M_j = X_1^{m_1} \cdots X_j^{m_j}$ (avec $M_0 = \mathrm{I}_2$). Remarquons que M_j est à coefficients entiers, pour tout $j \in [\![0, r]\!]$, puisque les matrices A, A^{-1}, B et B^{-1} sont à coefficients entiers. On note alors $c_j \in \mathbb{Z}$ le coefficient de M_j en haut à gauche si j est pair, respectivement le coefficient de M_j en haut à droite si j est impair.
On a donc $c_0 = 1$ et $c_1 = 2$. Nous allons montrer par récurrence que l'on a $|c_j| \geqslant j + 1$, pour tout $j \in [\![0, r]\!]$, ce qui prouvera en particulier que $|c_r| \geqslant r + 1$, donc que $M \neq \mathrm{I}_2$, et la liberté de la famille (A, B).

Remarquons que $M_{2k} = M_{2k-1}B^{m_{2k}}$ et que $M_{2k+1} = M_{2k}A^{m_{2k+1}}$, pour tout $k \in \left[\!\left[1, E\left(\frac{r}{2}\right)\right]\!\right]$, puisque $X_1 = A$.

Il est facile de voir que pour tout $m \in \mathbb{Z}$, on a

$$A^m = \begin{pmatrix} 1 & 2m \\ 0 & 1 \end{pmatrix} \quad \text{et} \quad B^m = \begin{pmatrix} 1 & 0 \\ 2m & 1 \end{pmatrix}.$$

On a donc

$$M_{2k} = \begin{pmatrix} c_{2k} & * \\ * & * \end{pmatrix} = \begin{pmatrix} * & c_{2k-1} \\ * & * \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 2m_{2k} & 1 \end{pmatrix} = \begin{pmatrix} * & c_{2k-1} \\ * & * \end{pmatrix}.$$

Ainsi, en identifiant les coefficients, on voit que

$$M_{2k} = \begin{pmatrix} c_{2k} & c_{2k-1} \\ * & * \end{pmatrix}.$$

De la même manière, on voit que

$$M_{2k+1} = \begin{pmatrix} c_{2k} & c_{2k+1} \\ * & * \end{pmatrix}.$$

En utilisant les relations $M_{2k} = M_{2k-1}B^{m_{2k}}$ et $M_{2k+1} = M_{2k}A^{m_{2k+1}}$, on obtient les égalités

$$c_{2k} = c_{2k-2} + 2m_{2k}c_{2k-1} \quad \text{et} \quad c_{2k+1} = c_{2k-1} + 2m_{2k+1}c_{2k}.$$

On a donc

$$|c_{2k}| \geqslant 2|m_{2k}| \times |c_{2k-1}| - |c_{2k-2}| \geqslant 2|c_{2k-1}| - |c_{2k-2}|,$$

car $m_{2k} \neq 0$. Ainsi, on a

$$|c_{2k}| - |c_{2k-1}| \geqslant |c_{2k-1}| - |c_{2k-2}|.$$

De même, on a

$$|c_{2k+1}| - |c_{2k}| \geqslant |c_{2k}| - |c_{2k-1}|.$$

Bref, pour tout $j \in [\![0, r]\!]$, on obtient

$$|c_{j+1}| - |c_j| \geqslant |c_j| - |c_{j-1}|.$$

Comme $|c_1| - |c_0| = 1$, une récurrence facile montre que $|c_j| \geqslant j + 1$, pour tout $j \in [\![0, r]\!]$, d'où le résultat.

En particulier, on peut construire un groupe libre sur un ensemble X à deux éléments, en prenant $F = \langle A, B \rangle$ et en considérant l'application injective $\iota : X \hookrightarrow F$ qui envoie les deux éléments de X sur A et B respectivement.

De même que toute application linéaire entre deux espaces vectoriels est uniquement déterminée par les images des éléments d'une base de l'espace de départ, il en est de même dans le cas des groupes.

4.6. Proposition. *Soit F un groupe, et soit $\iota : X \hookrightarrow F$ une application injective. Alors, (F, ι) est un groupe libre sur X si, et seulement si, (F, ι) satisfait la propriété universelle suivante : pour tout groupe F', et toute famille $\mathbf{a} = (a_x)_{x \in X}$ d'éléments de F' indexée par X, il existe un unique morphisme de groupes $\varphi_{\mathbf{a}} : F \longrightarrow F'$ tel que*

$$\varphi_{\mathbf{a}}\big(\iota(x)\big) = a_x, \quad \text{pour tout } x \in X.$$

Démonstration. Quitte à remplacer X par $\iota(X)$, on peut supposer que ι est l'inclusion, et donc que $X \subset F$.

Supposons que (F, ι) soit libre sur X.

D'après le lemme 4.2, tout élément $g \in F$ peut s'écrire de manière unique sous la forme

$$g = x_1^{n_1} \cdots x_k^{n_k}, \quad k \geqslant 0,\ n_i \in \mathbb{Z} \setminus \{0\},\ x_i \in X,$$

avec $x_i \neq x_{i+1}$, pour tout $i \in [\![1, k-1]\!]$. Si un morphisme $\varphi_{\mathbf{a}} : F \longrightarrow F'$ vérifiant les conditions de l'énoncé existe, on a nécessairement

$$\varphi_{\mathbf{a}}(x_1^{n_1} \cdots x_k^{n_k}) = \varphi_{\mathbf{a}}(x_1)^{n_1} \cdots \varphi_{\mathbf{a}}(x_k)^{n_k} = a_{x_1}^{n_1} \cdots a_{x_k}^{n_k},$$

ce qui démontre l'unicité. Il reste à voir que l'application $\varphi_{\mathbf{a}}$ définie par la formule précédente définit bien un morphisme de groupes.

Soient maintenant x, $y \in F$. On peut écrire

$$x = x_1^{n_1} \cdots x_r^{n_r} \quad \text{et} \quad y = x_{r+1}^{n_{r+1}} \cdots x_s^{n_s},$$

avec $x_1, \ldots, x_s \in X$, $n_i \in \mathbb{Z} \setminus \{0\}$, et $x_i \neq x_{i+1}$, pour tout $i \in [\![1, s]\!]$, sauf éventuellement si $i = r$.

Supposons tout d'abord que $x_r \neq x_{r+1}$. L'écriture de xy dans la base X est alors $xy = x_1^{n_1} \cdots x_s^{n_s}$.

Par définition de $\varphi_{\mathbf{a}}$, on a alors

$$\varphi_{\mathbf{a}}(xy) = \varphi_{\mathbf{a}}(x_1^{n_1} \cdots x_s^{n_s}) = a_{x_1}^{n_1} \cdots a_{x_s}^{n_s},$$

d'une part, et

$$\begin{aligned} \varphi_{\mathbf{a}}(x)\varphi_{\mathbf{a}}(y) &= \varphi_{\mathbf{a}}(x_1^{n_1} \cdots x_r^{n_r})\varphi_{\mathbf{a}}(x_{r+1}^{n_{r+1}} \cdots x_s^{n_s}) \\ &= (a_{x_1}^{n_1} \cdots a_{x_r}^{n_r})(a_{x_{r+1}}^{n_{r+1}} \cdots a_{x_s}^{n_s}) \\ &= a_{x_1}^{n_1} \cdots a_{x_s}^{n_s} \end{aligned}$$

d'autre part, d'où l'égalité $\varphi_{\mathbf{a}}(xy) = \varphi_{\mathbf{a}}(x)\varphi_{\mathbf{a}}(y)$.

Supposons maintenant que $x_r = x_{r+1}$. On a alors $x_r \neq x_{r+2}$, et l'écriture de xy dans la base X est

$$xy = x_1^{n_1} \cdots x_r^{n_r+n_{r+1}} x_{r+2}^{n_{r+2}} \cdots x_s^{n_s}.$$

Ainsi, on a

$$\varphi_{\mathbf{a}}(xy) = a_{x_1}^{n_1} \cdots a_{x_r}^{n_r+n_{r+1}} a_{x_{r+2}}^{n_{r+2}} \cdots a_{x_s}^{n_s}.$$

Mais, comme $x_r = x_{r+1}$, on a $a_{x_r} = a_{x_{r+1}}$, et l'on a aussi

$$\begin{aligned} \varphi_{\mathbf{a}}(x)\varphi_{\mathbf{a}}(y) &= (a_{x_1}^{n_1} \cdots a_{x_r}^{n_r})(a_{x_{r+1}}^{n_{r+1}} \cdots a_{x_s}^{n_s}) \\ &= a_{x_1}^{n_1} \cdots a_{x_r}^{n_r+n_{r+1}} a_{x_{r+2}}^{n_{r+2}} \cdots a_{x_s}^{n_s} \\ &= \varphi_{\mathbf{a}}(xy). \end{aligned}$$

Ainsi, $\varphi_{\mathbf{a}}$ est bien un morphisme de groupes.

Inversement, supposons maintenant que pour tout groupe F', et toute famille $\mathbf{a} = (a_x)_{x\in X}$ d'éléments de F' indexée par X, il existe un unique morphisme de groupes $\varphi_{\mathbf{a}} : F \longrightarrow F'$ tel que

$$\varphi_{\mathbf{a}}(x) = a_x, \quad \text{pour tout } x \in X.$$

En particulier, il existe un unique morphisme de groupes $\varphi : F \longrightarrow \langle X\rangle$ tel que

$$\varphi(x) = x, \quad \text{pour tout } x \in X.$$

Notons que φ est surjective. Le morphisme $\begin{array}{rcl} \psi : F & \longrightarrow & F \\ g & \longmapsto & \varphi(g) \end{array}$ vérifie alors également

$$\psi(x) = x, \quad \text{pour tout } x \in X.$$

Mais, Id_F est un morphisme de groupes vérifiant la même propriété. Par unicité, on a alors $\psi = \mathrm{Id}_F$.

En comparant les images de ces morphismes, on obtient $\langle X\rangle = F$. Ainsi, la partie X est génératrice. Il reste à montrer que X est libre.

Soit $x \in F$ un élément de la forme $x = x_1^{n_1} \cdots x_k^{n_k}$, avec $k \geqslant 1, n_i \in \mathbb{Z}\setminus\{0\}$, et $x_i \in X$ vérifiant $x_i \neq x_{i+1}$ pour tout $i \in [\![1, k-1]\!]$.

Soient $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$ et $B = \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix}$. Par hypothèse, il existe un morphisme de groupes $\varphi : F \longrightarrow \langle A, B\rangle$ tel que $\varphi(x_j) = A$ si j est pair, et $\varphi(x_j) = B$ si j est impair. Puisque l'on a $\varphi(x_i) \neq \varphi(x_{i+1})$ pour tout $i \in [\![1, k-1]\!]$, et puisque $\{A, B\}$ est une base de $\langle A, B\rangle$ par l'exemple 4.5 (4), on obtient

$$\varphi(x_1)^{n_1} \cdots \varphi(x_k)^{n_k} \neq \mathrm{I}_2,$$

soit encore $\varphi(x) \neq \varphi(1_F)$. En particulier, $x \neq 1_F$, ce qu'il fallait vérifier. Autrement dit, X est libre. C'est donc une base, et cela achève la démonstration. □

4.7. Définition. Soit (F, ι) un groupe libre sur un ensemble X (non vide). Si G est un groupe, et si $\mathbf{a} = (a_x)_{x\in X}$ est une famille d'éléments de G indexée par X, l'unique morphisme de groupes $\varphi_{\mathbf{a}} : F \longrightarrow G$ tel que

$$\varphi_{\mathbf{a}}\big(\iota(x)\big) = a_x, \quad \text{pour tout } x \in X$$

est appelé le *morphisme d'évaluation en* $\mathbf{a}$.

Pour tout $g \in F$, l'image de g par $\varphi_{\mathbf{a}}$ est notée parfois $g(\mathbf{a})$.
Si l'on écrit

$$g = \iota(x_1)^{\varepsilon_1} \cdots \iota(x_k)^{\varepsilon_k}, \quad x_i \in X,\ \varepsilon_i \in \{\pm 1\},$$

ce qui est légitime puisque $\iota(X)$ engendre[6] F, alors on aura

$$\varphi_{\mathbf{a}}(g) = a_{x_1}^{\varepsilon_1} \cdots a_{x_k}^{\varepsilon_k}.$$

Autrement dit, on a « évalué » le générateur $\iota(x)$ en a_x, pour tout $x \in X$.

4.8. Remarque. La plupart des ouvrages définissent un groupe libre sur un ensemble X comme étant un couple (F, ι) satisfaisant à la propriété universelle décrite dans la proposition précédente. Néanmoins, nous avons préféré une approche en termes d'existence d'une base, plus concrète, et qui fait bien ressentir la signification du terme « libre ». En revanche, pour les démonstrations ou les calculs, la propriété la plus utile est la propriété universelle d'un groupe libre de base X.

Nous allons maintenant nous intéresser à l'existence d'un groupe libre sur un ensemble X arbitraire. L'existence d'un tel groupe n'est pas immédiate. En revanche, s'il existe, il est unique, à unique isomorphisme près.

4.9. Lemme. *Pour tous groupes libres (F, ι) et (F', ι') de base X, il existe un unique isomorphisme*

$$\varphi : (F, \iota) \xrightarrow{\sim} (F', \iota').$$

C'est l'unique morphisme de groupes $\varphi : F \longrightarrow F'$ tel que

$$\varphi\big(\iota(x)\big) = \iota'(x), \quad \textit{pour tout } x \in X.$$

Démonstration. Si un tel isomorphisme de groupes libres existe, il vérifie la condition de l'énoncé par la proposition 4.6. En appliquant la propriété universelle de (F, ι) à l'application ι', on constate qu'un tel morphisme de groupes $\varphi : F \longrightarrow F'$ existe et est unique. Il reste donc à vérifier que c'est un isomorphisme. En appliquant la propriété universelle de (F', ι') à l'application ι, on obtient un morphisme de groupes $\psi : F' \longrightarrow F$ tel que

$$\psi\big(\iota'(x)\big) = \iota(x), \quad \text{pour tout } x \in X.$$

Remarquons maintenant que $\psi \circ \varphi : F \longrightarrow F$ est un morphisme de groupes

6. On peut même rendre l'écriture unique en simplifiant et en regroupant des termes, mais ce n'est pas utile ici.

vérifiant

$$(\psi \circ \varphi)\big(\iota(x)\big) = \iota(x), \quad \text{pour tout } x \in X.$$

Mais, $\mathrm{Id}_F : F \longrightarrow F$ est un autre morphisme vérifiant cette propriété. Puisque (F, ι) est un groupe libre de base X, on en déduit que $\psi \circ \varphi = \mathrm{Id}_F$. On démontre de même que $\varphi \circ \psi = \mathrm{Id}_{F'}$. Ainsi, φ est un isomorphisme, d'où la conclusion. □

Avant de démontrer l'existence d'un groupe libre de base X, montrons comment un tel groupe permet de résoudre les questions évoquées en début de paragraphe. On commence par définir ce qu'est une relation.

4.10. Définition. Soit (F, ι) un groupe libre de base X. Soit $\mathbf{a} = (a_x)_{x \in X}$ une famille d'éléments d'un groupe G. Soit $R \subset F$. On dit que $\mathbf{a}$ *vérifie les relations* R si $r(\mathbf{a}) = 1_G$ pour tout $r \in R$.

On dit que G vérifie les relations R, s'il existe au moins une famille d'éléments de G vérifiant les relations R.

4.11. Exemple

Soit (F, ι) un groupe libre de base $X = \{a, b\}$, avec $|X| = 2$ (par exemple, on peut prendre le groupe libre de l'exemple 4.5 (4)).

Alors, les groupes D_4 et $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ vérifient les relations a^4, b^2, $bab^{-1}a$. Autrement dit, dans chacun de ces deux groupes, il existe deux éléments σ et τ vérifiant

$$\sigma^4 = 1, \quad \tau^2 = 1, \quad \tau\sigma\tau^{-1} = \sigma^{-1}.$$

4.12. Définition. Soit (F, ι) un groupe libre de base X. Un groupe *présenté par les générateurs* X *et les relations* R est un couple (G, η), où G est un groupe et $\eta : X \hookrightarrow G$ est une application injective vérifiant les conditions suivantes :

(1) le groupe G est engendré par $\eta(X)$;

(2) la famille $\big(\eta(x)\big)_{x \in X}$ vérifie les relations R ;

(3) pour tout groupe G', et toute famille $\mathbf{a} = (a_x)_{x \in X}$ d'éléments de G' vérifiant les relations R, il existe un unique morphisme $\varphi : G \longrightarrow G'$ tel que

$$\varphi\big(\eta(x)\big) = a_x, \quad \text{pour tout } x \in X.$$

On dit que deux groupes (G, η) et (G', η') présenté par les générateurs X et les relations R sont *isomorphes* s'il existe un isomorphisme $\varphi : G \xrightarrow{\sim} G'$ tel que

$$\varphi\big(\eta(x)\big) = \eta'(x), \quad \text{pour tout } x \in X.$$

4.13. Remarque

Un groupe présenté par les générateurs X et les relations R est donc la donnée d'un groupe G vérifiant les relations R, et d'une famille d'éléments spécifiques vérifiant ces relations. Un isomorphisme entre deux tels groupes est alors un isomorphisme de groupes envoyant la famille choisie dans le premier groupe sur la famille choisie dans le second groupe.

Souvent, on ne spécifie pas η.

La démonstration du lemme suivant est identique en tout point à celle du lemme 4.9, et est laissée au lecteur.

4.14. Lemme. *Pour tous groupes (G,η) et (G',η') présentés par les générateurs X et les relations R, il existe un unique isomorphisme*

$$\varphi : (G,\eta) \xrightarrow{\sim} (G',\eta').$$

C'est l'unique morphisme de groupes $\varphi : G \longrightarrow G'$ tel que

$$\varphi\big(\eta(x)\big) = \eta'(x), \qquad \textit{pour tout } x \in X.$$

4.15. Remarque

Si (F,ι) et (F',ι') sont deux groupes libres sur un même ensemble X, et si $\theta : (F,\iota) \xrightarrow{\sim} (F',\iota')$ est un isomorphisme, alors (G,η) est présenté par les générateurs X et les relations R si, et seulement si, (G,η) est présenté par les générateurs X et les relations $\theta(R)$.

En effet, notons $\varphi_\eta : F \longrightarrow G$ le morphisme d'évaluation en η relatif au groupe F, et $\varphi'_\eta : F' \longrightarrow G$ le morphisme d'évaluation en η relatif au groupe F'.

On a alors

$$(\varphi_\eta \circ \theta^{-1})\big(\iota'(x)\big) = (\varphi_\eta \circ \theta^{-1})\Big(\theta\big(\iota(x)\big)\Big) = \varphi_\eta\big(\iota(x)\big) = \eta_x, \quad \text{pour tout } x \in X.$$

Ainsi, on a $\varphi'_\eta = \varphi_\eta \circ \theta^{-1}$. Le reste découle immédiatement de cette égalité.

Décrivons un peu plus en détail ce que cela signifie. Pour tout $x \in X$, on pose $y_x = \iota(x)$ et $z_x = \iota'(x)$. Ainsi, $(y_x)_{x\in X}$ est une base de F et $(z_x)_{x\in X}$ est une base de F'. L'isomorphisme $\theta : (F,\iota) \xrightarrow{\sim} (F',\iota')$ envoie alors y_x sur z_x pour tout $x \in X$.

Si $r \in R$, on peut alors écrire $r = y_{x_1}^{\varepsilon_1} \cdots y_{x_k}^{\varepsilon_k}$, avec $x_i \in X, \varepsilon_i \in \{\pm 1\}$. Le morphisme d'évaluation en η envoie par définition y_x sur η_x, et la relation r sera vérifiée par η si, et seulement si, on a

$$\eta_{x_1}^{\varepsilon_1} \cdots \eta_{x_k}^{\varepsilon_k} = 1_G.$$

Par ailleurs, l'isomorphisme θ envoie y_x sur z_x, donc $\theta(r) = z_{x_1}^{\varepsilon_1} \cdots z_{x_k}^{\varepsilon_k}$, et la relation $\theta(r)$ sera vérifiée par η si, et seulement si, on a

$$\eta_{x_1}^{\varepsilon_1} \cdots \eta_{x_k}^{\varepsilon_k} = 1_G.$$

En particulier, cela montre que la notion de groupe libre présenté par des générateurs X et des relations R ne dépend aucunement du groupe libre de base X dans lequel on a choisi les relations. Intuitivement, cela se comprend bien. Les relations satisfaites dans un groupe G ne font pas apparaître spécifiquement les éléments du groupe libre. Le groupe libre est simplement utilisé pour formaliser la notion de relation.

C'est pourquoi bien souvent, on ne précise pas dans quel groupe libre sur X sont prises les relations R.

Nous allons maintenant démontrer (modulo l'existence d'un groupe libre sur X) que pour tout ensemble X et tout ensemble de relations R, il existe au moins un groupe présenté par les générateurs X et les relations R.

4.16. Proposition. *Soit X un ensemble, soit (F, ι) un groupe libre sur X et soit $R \subset F$. Soit H_R l'intersection des sous-groupes distingués de F contenant R, et soit*

$$\begin{aligned} \pi : F &\longrightarrow F/H_R \\ y &\longmapsto \overline{y} \end{aligned}$$

la projection canonique. Posons $\eta_R = \big(\overline{\iota(x)}\big)_{x\in X}$. Alors, H_R est un sous-groupe distingué de F contenant R, et le couple $(F/H_R, \eta_R)$ est un groupe présenté par les générateurs X et les relations R.

Démonstration. Puisque $\iota(X) = \big(\iota(x)\big)_{x\in X}$ engendre le groupe F, son image par π, soit $\eta_R = \big(\overline{\iota(x)}\big)_{x\in X}$, engendre $\pi(F) = F/H_R$.

De plus, il est clair que l'intersection d'une famille de sous-groupes distingués de F contenant R est un sous-groupe distingué de F contenant R.

Si $r = \iota(x_1)^{\varepsilon_1} \cdots \iota(x_k)^{\varepsilon_k}$, avec $x_i \in X$, $\varepsilon_i \in \{\pm 1\}$, on a

$$\varphi_\eta(r) = \overline{\iota(x_1)}^{\varepsilon_1} \cdots \overline{\iota(x_k)}^{\varepsilon_k} = \overline{\iota(x_1)^{\varepsilon_1} \cdots \iota(x_k)^{\varepsilon_k}} = \overline{r} = \overline{1}_F,$$

car $r \in H_R$. Ainsi, η satisfait les relations R.

Soit G un groupe possédant une famille d'éléments $\mathbf{a} = (a_x)_{x\in X}$ vérifiant les relations R. Supposons qu'il existe un morphisme $\varphi : F/H_R \longrightarrow G'$ tel que

$$\varphi\big(\overline{\iota(x)}\big) = a_x, \qquad \text{pour tout } x \in X.$$

Alors, φ est unique. En effet, tout élément de F/H_R est produit d'éléments de η_R et de leurs inverses, et donc les relations précédentes déterminent complètement l'image de chaque élément de F/H_R.

Montrons maintenant l'existence d'un tel morphisme. Par choix de $\mathbf{a}$, on a l'inclusion $R \subset \mathrm{Ker}(\varphi_{\mathbf{a}})$. Comme $\mathrm{Ker}(\varphi_{\mathbf{a}})$ est un sous-groupe distingué de F contenant R, on a $H_R \subset \mathrm{Ker}(\varphi_{\mathbf{a}})$ par définition. Par le théorème de factorisation, $\varphi_{\mathbf{a}}$ induit alors par passage au quotient un morphisme de groupes $\overline{\varphi}_{\mathbf{a}} : F/H_R \longrightarrow G$ tel que $\varphi_{\mathbf{a}} = \overline{\varphi}_{\mathbf{a}} \circ \pi$. En particulier, on a

$$\overline{\varphi}_{\mathbf{a}}\big(\overline{\iota(x)}\big) = \varphi_{\mathbf{a}}\big(\iota(x)\big) = a_x, \quad \text{pour tout } x \in X.$$

Cela achève la démonstration. □

4.17. Définition. L'unique groupe (à isomorphisme près) présenté par les générateurs X et les relations R est noté $\langle X \mid R\rangle$.

D'après la proposition 4.16, si (F, ι) est un groupe libre sur X, et si H_R est le plus petit sous-groupe distingué de F contenant les relations R, alors $\big(F/H_R, \big(\overline{\iota(x)}\big)_{x\in X}\big)$ est un tel groupe.

L'unicité à isomorphisme près montre que pour tout groupe (G, η) présenté par les générateurs X et les relations R, il existe un isomorphisme

$$G \simeq \langle X \mid R\rangle.$$

On dit alors aussi que $\langle X \mid R\rangle$ est une *présentation* du groupe G.

4.18. Remarque. En général, on fait l'abus de notation suivant : on note le groupe $\langle X \mid R\rangle$ sous la forme $\langle x \in X \mid r = 1,\ r \in R\rangle$.

On identifie aussi souvent X avec son image $\eta(X)$.

4.19. Exemples

(1) Les considérations faites en début de paragraphe montre que l'on a un isomorphisme
$$D_4 \simeq \langle a, b \mid a^4 = b^2 = 1,\ bab^{-1} = a^{-1}\rangle.$$

(2) Soit $n \geqslant 1$. Alors, $\langle a \mid a^n = 1\rangle \simeq \mathbb{Z}/n\mathbb{Z}$.
En effet, l'élément $\overline{1}$ engendre $\mathbb{Z}/n\mathbb{Z}$, et vérifie $n\cdot\overline{1} = \overline{0}$. Si maintenant G est un groupe contenant un élément $g \in G$ vérifiant $g^n = 1_G$, alors le morphisme
$$\begin{aligned} h_g\colon \mathbb{Z} &\longrightarrow G \\ m &\longmapsto g^m \end{aligned}$$
induit par passage au quotient un morphisme $\overline{h}_g : \mathbb{Z}/n\mathbb{Z} \longrightarrow G$ qui envoie $\overline{1}$ sur g.

(3) Soit $n \geqslant 3$. Alors, $G = \langle a, b \mid a^n = 1,\ b^2 = 1,\ bab^{-1} = a^{-1}\rangle \simeq D_n$.
En effet, en utilisant le fait que tout élément de G est produit d'éléments a, a^{-1}, b ou b^{-1}, ainsi que les relations qu'ils vérifient dans G,

on voit que tout élément de F/H_R peut s'écrire sous la forme

$$\overline{a}^i\overline{b}^j, \quad i \in [\![0, n-1]\!], \ j = 0, 1.$$

En particulier, $|G| \leqslant 2n$.

Soient $\sigma, \tau \in D_n$ la rotation d'angle de mesure $\frac{2\pi}{n}$ et la symétrie orthogonale par rapport à l'axe (Ox), respectivement. Comme σ et τ vérifient les relations

$$\sigma^n = \mathrm{Id}, \quad \tau^2 = \mathrm{Id}, \quad \tau\sigma\tau^{-1} = \sigma^{-1},$$

il existe un unique morphisme de groupes $\varphi : G \longrightarrow D_n$ qui envoie $\overline{a}$ sur σ, et $\overline{b}$ sur τ. Ce morphisme est surjectif, car σ et τ engendrent le groupe D_n. En particulier, $|G| \geqslant 2n$. Ainsi, $|G| = 2n$, et φ est donc un isomorphisme.

Le lemme suivant montre que cette notion est non sans intérêt.

4.20. Lemme. *Tout groupe G admet au moins une présentation par générateurs et relations.*

Démonstration. Soit G un groupe, et soit (F, ι) un groupe libre sur G. Considérons alors l'unique morphisme $\varphi : F \longrightarrow G$ tel que

$$\varphi\big(\iota(g)\big) = g, \quad \text{pour tout } g \in G.$$

Ce morphisme est par définition surjectif. Posons $R = \mathrm{Ker}(\varphi)$. Alors, l'intersection H_R de tous les sous-groupes distingués contenant R est simplement $\mathrm{Ker}(\varphi)$. Par la proposition 4.16, $\big(F/\mathrm{Ker}(\varphi), \big(\overline{\iota(g)}\big)_{g\in G}\big)$ est donc un groupe présenté par les générateurs G et les relations R. De plus, on a un isomorphisme $\overline{\varphi} : G/\mathrm{Ker}(\varphi) \xrightarrow{\sim} G$, ce qui achève la démonstration. □

4.21. Définition. Soit G un groupe. On dit que G est de *présentation finie* s'il admet une présentation $\langle X \mid R\rangle$, avec X et R finis.

4.22. Remarque. Un groupe de présentation finie est nécessairement de type fini, i.e. engendré par un nombre fini d'éléments, et cela par définition. En revanche, il existe des contre-exemples à la réciproque.

4.23. Lemme. *Tout groupe fini est de présentation finie.*

Démonstration. Soit G un groupe fini. Soit (F, ι) un groupe libre sur G, et soit $R = \big\{\iota(g_1)\iota(g_2)\iota(g_1g_2)^{-1} \mid g_1, g_2 \in G\big\}$. Comme G est fini, R est aussi fini.

Soit $\varphi : F \longrightarrow G$ l'unique morphisme tel que

$$\varphi\big(\iota(g)\big) = g, \quad \text{pour tout } g \in G.$$

Ce morphisme est clairement surjectif.

De plus, $R \subset \mathrm{Ker}(\varphi)$ puisque, pour tous g_1, $g_2 \in G$, on a
$\varphi\big(\iota(g_1)\iota(g_2)\iota(g_1g_2)^{-1}\big) = \varphi\big(\iota(g_1)\big)\varphi\big(\iota(g_2)\big)\varphi\big(\iota(g_1g_2)\big)^{-1} = g_1g_2(g_1g_2)^{-1} = 1_G$.
Ainsi, $\varphi : F \longrightarrow G$ induit par passage au quotient un morphisme surjectif

$$\overline{\varphi} : F/H_R \longrightarrow G.$$

En particulier, on a $|F/H_R| \geqslant |G|$. Remarquons maintenant que l'application

$$\begin{array}{rcl} \psi : G & \longrightarrow & F/H_R \\ g & \longmapsto & \overline{\iota(g)} \end{array}$$

est un morphisme de groupes. En effet, puisque $R \subset H_R$, on a

$$\begin{array}{rcl} \psi(g_1)\psi(g_2) & = & \overline{\iota(g_1)}\,\overline{\iota(g_2)} \;=\; \overline{\iota(g_1)\iota(g_2)} \\ & = & \overline{(\iota(g_1)\iota(g_2)\iota(g_1g_2)^{-1})\iota(g_1g_2)} \\ & = & \overline{\iota(g_1)\iota(g_2)\iota(g_1g_2)^{-1}}\;\overline{\iota(g_1g_2)} \\ & = & \overline{\iota(g_1g_2)} \;=\; \psi(g_1g_2), \end{array}$$

pour tous g_1, $g_2 \in G$. Puisque la famille $\big(\overline{\iota(g)}\big)_{g\in G}$ engendre F/H_R, on en déduit que ψ est surjective. Par conséquent, $|G| \geqslant |F/H_R|$, et l'on obtient finalement

$$|G| = |F/H_R|.$$

Ainsi, $\overline{\varphi}$ est un isomorphisme. Comme $\big(F/H_R, \big(\overline{\iota(g)}\big)_{g\in G}\big)$ est un groupe présenté par les générateurs G et les relations R, cela achève la démonstration. □

Nous allons maintenant montrer comment les présentations de groupes permettent de déterminer, au moins en théorie, tous les morphismes d'un groupe G donné dans un groupe G' arbitraire.

4.24. Proposition. *Soit (G, η) un groupe présenté par les générateurs X et les relations R. Alors, pour tout groupe G', l'ensemble des morphismes de groupes $G \longrightarrow G'$ est en bijection avec l'ensemble des familles $\mathbf{a} = (a_x)_{x\in X}$ d'éléments de G' vérifiant les relations R.*

Plus précisément la correspondance est donnée par

$$\begin{array}{rcl} f & \longmapsto & \big(f(\eta_x)\big)_{x\in X} \\ f_{\mathbf{a}} & \longmapsfrom & \mathbf{a}, \end{array}$$

où $f_{\mathbf{a}} : G \longrightarrow G'$ est l'unique morphisme tel que

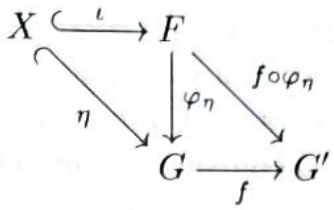
$$f_{\mathbf{a}}\big(\eta(x)\big) = a_x, \quad \text{pour tout } x \in X.$$

Démonstration. Soit (F, ι) un groupe libre sur l'ensemble X tel que $R \subset F$, et soit $\eta = (\eta_x)_{x\in X}$ une famille génératrice de G vérifiant les relations R. Alors, pour tout morphisme $f : G \longrightarrow G'$, la famille $\mathbf{a}_f = \big(f(\eta_x)\big)_{x\in X}$ vérifie les relations R.

En effet, si $\varphi_\eta : F \longrightarrow G$ est le morphisme d'évaluation en η, alors

$$f \circ \varphi_\eta : F \longrightarrow G'$$

est un morphisme qui envoie $\iota(x)$ sur $f(\eta_x)$ pour tout $x \in X$.



C'est donc le morphisme d'évaluation en $\mathbf{a}_f$. Pour tout $r \in R$, on a alors

$$\varphi_{\mathbf{a}_f}(r) = f\big(\varphi_\eta(r)\big) = f(1_G) = 1_{G'}.$$

Si maintenant $\mathbf{a}$ est une famille d'éléments de G vérifiant les relations R, par définition de G, on a un unique morphisme $f_{\mathbf{a}} : G \longrightarrow G'$ tel que

$$f_{\mathbf{a}}(\eta_x) = a_x, \qquad \text{pour tout } x \in X.$$

Il reste à voir que ces deux constructions sont inverses l'une de l'autre.

Si $\mathbf{a}$ est une famille d'éléments de G vérifiant les relations R, alors on a

$$\mathbf{a}_{f_{\mathbf{a}}} = \big(f_{\mathbf{a}}(\eta_x)\big)_{x\in X} = (a_x)_{x\in X} = \mathbf{a}.$$

Inversement, si $f : G \longrightarrow G'$ est un morphisme de groupes, $f_{\mathbf{a}_f}$ est l'unique morphisme $G \longrightarrow G'$ qui envoie η_x sur $f(\eta_x)$ pour tout $x \in X$. Mais f est un tel morphisme, d'où l'égalité $f_{\mathbf{a}_f} = f$. Cela achève la démonstration. □

Il reste maintenant à construire un groupe libre sur X, pour tout ensemble X. Pour cela, on va partir d'une idée très simple : formellement, un élément d'un groupe libre G de base X est une suite finie de caractères $x_1^{\varepsilon_1} \cdots x_k^{\varepsilon_k}$, $x_i \in X$, $\varepsilon_i \in \{\pm 1\}$, et les seules relations possibles entre les éléments de G sont les relations évidentes

$$gg^{-1} = g^{-1}g = 1, \quad g \in G.$$

4.25. Définition. Soit X un ensemble. Pour chaque $x \in X$, on introduit un nouveau symbole x^*, et l'on note $X^* = \{x^* \mid x \in X\}$. C'est un ensemble équipotent à X, et disjoint[7] de X.

Un *mot* sur X est une suite finie $\omega = (w_1, \ldots, w_n)$ d'éléments de $X \cup X^*$. L'entier n s'appelle la *longueur* de ω. Si $n = 0$, on obtient le *mot vide* $\emptyset$. C'est l'unique mot de longueur 0.

En général, on note plutôt $\omega = w_1 \cdots w_n$. L'ensemble des mots sur X sera noté $\mathcal{M}_X$. L'ensemble $\mathcal{M}_X$ est alors muni d'une loi interne, dite loi de *concaténation* définie par

$$\begin{aligned} \mathcal{M}_X \times \mathcal{M}_X &\longrightarrow \mathcal{M}_X \\ (w_1 \cdots w_n, w'_1 \cdots w'_m) &\longmapsto w_1 \cdots w_m w'_1 \cdots w'_m. \end{aligned}$$

En particulier, $\omega\emptyset = \emptyset\omega = \omega$ pour tout $\omega \in \mathcal{M}_X$.

La loi de concaténation est associative, d'élément neutre $\emptyset$.

7. Sauf si l'on n'a vraiment pas de chance
Dans ce cas, on remplace le symbole $*$ par autre chose.

Bien sûr, la loi de concaténation sur $\mathcal{M}_X$ n'est pas une loi de groupe, aucun mot non vide n'ayant de symétrique. L'idée est maintenant de forcer les éléments de X à posséder un inverse. Pour cela, nous allons définir une relation d'équivalence sur $\mathcal{M}_X$.

Avant cela, nous allons introduire la notion de sous-mot, qui sera bien pratique pour les démonstrations.

4.26. Définition. Soient ω, ω' deux mots sur X. On dit que ω' est un *sous-mot* de ω s'il existe deux mots ω_1 et ω_2 (éventuellement vides) tels que

$$\omega = \omega_1 \omega' \omega_2.$$

On dit que le mot ω *débute* par ω' s'il existe un mot ω'' (éventuellement vide) tel que $\omega = \omega'\omega''$. On dit que le mot ω *se termine* par ω' s'il existe un mot ω'' (éventuellement vide) tel que $\omega = \omega''\omega'$.

S'il existe des mots ω_1, ω_2, ω'' tels que $\omega = \omega_1\omega_2$ et $\omega' = \omega_1\omega''\omega_2$, on dit que l'on a obtenu ω' à partir de ω en *insérant* ω''. S'il existe des mots ω_1, ω_2, ω'' tels que $\omega = \omega_1\omega''\omega_2$ et $\omega' = \omega_1\omega_2$, on dit que l'on a obtenu ω' à partir ω en *effaçant* ω''.

Nous allons maintenant définir notre relation d'équivalence sur $\mathcal{M}_X$.

4.27. Définition. Deux mots ω et ω' seront dit *élémentairement équivalents* si $\omega = \omega'$, ou si ω' peut se déduire de ω en insérant ou en effaçant un sous-mot de la forme xx^* ou x^*x, $x \in X$. On note cela $\omega \sim_{X,e} \omega'$.

Deux mots ω et ω' seront dit *équivalents* s'il existe une suite finie de mots

$$\omega_0 = \omega,\ \omega_1,\ \ldots,\ \omega_{r-1},\ \omega_r = \omega'$$

tels que les mots ω_i et ω_{i+1} soient élémentairement équivalents pour tout entier $i \in [\![0, r-1]\!]$. On note cela $\omega \sim_X \omega'$.

4.28. Lemme. *La relation $\sim_X$ est une relation d'équivalence sur l'ensemble des mots $\mathcal{M}_X$. De plus, pour tous mots ω, ω', ρ, ρ', on a*

$$\omega \sim_X \omega' \quad et \quad \rho \sim_X \rho' \quad \Longrightarrow \quad \omega\rho \sim_X \omega'\rho'.$$

Démonstration. On commence par remarquer que deux mots élémentairement équivalents sont équivalents, et que la relation d'équivalence élémentaire est réflexive et symétrique. En effet, la réflexivité est claire et la symétrie vient du fait que ω' se déduit de ω en insérant/effaçant un sous-mot de la forme xx^* ou x^*x, alors ω se déduit de ω' en effaçant/insérant ce même sous-mot.

On en déduit alors que la relation $\sim_X$ est réflexive et symétrique. Il est tout aussi immédiat à partir de la définition qu'elle est transitive. Montrons la deuxième assertion.

Remarquons tout d'abord que, vu la définition, si $\omega \sim_{X,e} \omega'$, alors pour tout $\rho \in \mathcal{M}_X$, on a

$$\rho\omega \sim_{X,e} \rho\omega' \quad \text{et} \quad \omega\rho \sim_{X,e} \omega'\rho.$$

En considérant une chaîne finie d'équivalences élémentaires

$$\omega = \omega_0 \sim_{X,e} \omega_1 \sim_{X,e} \cdots \sim_{X,e} \omega_{r-1} \sim_{X,e} \omega_r = \omega',$$

on en déduit que si $\omega \sim_X \omega'$, alors $\omega\rho \sim_X \omega'\rho$ pour tout $\rho \in \mathcal{M}_X$. Considérons maintenant deux chaîne finies d'équivalences élémentaires

$$\omega = \omega_0 \sim_{X,e} \omega_1 \sim_{X,e} \cdots \sim_{X,e} \omega_{r-1} \sim_{X,e} \omega_r = \omega',$$

et

$$\rho = \rho_0 \sim_{X,e} \rho_1 \sim_{X,e} \cdots \sim_{X,e} \rho_{s-1} \sim_{X,e} \rho_s = \rho'.$$

Quitte à rajouter des équivalences du type $\omega \sim_{X,e} \omega$ ou $\rho \sim_{X,e} \rho$, on peut supposer que $r = s$. Pour tout $i \in [\![0, r-1]\!]$, on a

$$\omega_i\rho_i \sim_{X,e} \omega_{i+1}\rho_i \sim_{X,e} \omega_{i+1}\rho_{i+1}.$$

Mais alors, on en déduit que $\omega_0\rho_0 \sim_X \omega_r\rho_r$, c'est-à-dire $\omega\rho \sim_X \omega'\rho'$. Cela achève la démonstration. □

Notation. On note F_X l'ensemble quotient $\mathcal{M}_X/\sim_X$.

4.29. Corollaire. *La loi interne*

$$F_X \times F_X \longrightarrow F_X$$
$$(\overline{\omega}_1, \overline{\omega}_2) \longmapsto \overline{\omega_1\omega_2}$$

est bien définie, et munit F_X d'une structure de groupe, de neutre $\overline{\emptyset}$.

De plus, pour tout $x \in X$, on a $\overline{x}^{-1} = \overline{x^*}$. Enfin, F_X est engendré par les éléments $\overline{x}, x \in X$.

Démonstration. Le lemme précédent assure que la loi de concaténation sur l'ensemble $\mathcal{M}_X$ passe au quotient, et induit donc une loi interne sur F_X. Puisque la loi de concaténation est associative, de neutre $\emptyset$, la loi sur F_X est associative, de neutre $\overline{\emptyset}$. De plus, par définition de la relation $\sim_X$, les mots xx^* et x^*x sont équivalents au mot vide $\emptyset$, et cela pour tout $x \in X$. On a donc

$$\overline{x}\ \overline{x^*} = \overline{xx^*} = \overline{\emptyset}, \quad \text{pour tout } x \in X,$$

et de même

$$\overline{x^*}\ \overline{x} = \overline{x^*x} = \overline{\emptyset}, \quad \text{pour tout } x \in X.$$

Ainsi, pour tout $x \in X$, $\overline{x}$ est inversible, d'inverse $\overline{x^*}$. Par suite, $\overline{x^*}$ est également inversible, d'inverse $\overline{x}$. Si maintenant $\omega = w_1 \cdots w_k$ est un mot sur X arbitraire, alors $\overline{\omega} = \overline{w}_1 \cdots \overline{w}_k$ est inversible, d'inverse $\overline{w}_k^{-1} \cdots \overline{w}_1^{-1}$.

Par conséquent, F_X est bien un groupe. Comme tout élément de F_X est par définition un produit de classes $\overline{x}, x \in X$, ou de leurs inverses $\overline{x^*}, x \in X$, le dernier point s'ensuit. □

Soit $\iota_X : X \longrightarrow F_X$ l'application $\begin{array}{rcl} \iota_X : X & \longrightarrow & F_X \\ x & \longmapsto & \overline{x}. \end{array}$

Notre but est maintenant de montrer que (F_X, ι_X) est un groupe libre sur X. Pour cela, nous avons besoin de la notion de mot réduit.

4.30. Définition. Un mot sur X est dit *réduit* s'il ne contient pas de sous-mot de la forme xx^* ou x^*x, avec $x \in X$.

Le mot vide est réduit, de même que tout mot de longueur 1.

4.31. Proposition. *Tout mot sur X est équivalent à un unique mot réduit. En particulier, deux mots réduits sont équivalents si, et seulement si, ils sont égaux.*

Démonstration. Afin d'alléger les écritures, nous allons introduire une notation. Si $w \in X \cup X^*$, on pose $w^* = x^*$ si $w = x$ pour un certain $x \in X$, et $w^* = x$ si $w = x^*$ pour un certain $x^* \in X^*$.

Soit $w \in X \cup X^*$. Pour tout $\rho \in \mathcal{M}_X$, on pose

$$T(w)(\rho) = \begin{cases} w\rho & \text{si } \rho \text{ ne débute pas par } w^* \\ \rho' & \text{si } \rho = w^*\rho'. \end{cases}$$

Si maintenant $\omega = w_1 \cdots w_k, w_i \in X \cup X^*$, on pose

$$T(\omega) = T(w_1) \circ \cdots \circ T(w_k).$$

En particulier, $T(\emptyset) = \mathrm{Id}_{\mathcal{M}_X}$.

On va montrer les propriétés suivantes. Soient $\omega, \omega', \rho \in \mathcal{M}_X$. Alors :

(1) si ρ est réduit, alors $T(\omega)(\rho)$ est réduit. En particulier, $T(\omega)(\emptyset)$ est réduit ;

(2) on a $T(\omega)(\rho) \sim_X \omega\rho$. En particulier, $T(\omega)(\emptyset) \sim_X \omega$;

(3) si ω est réduit, alors $T(\omega)(\emptyset) = \omega$;

(4) si $\omega \sim_X \omega'$, alors $T(\omega)(\emptyset) = T(\omega')(\emptyset)$.

(1) On commence par démontrer que pour tout $w \in X \cup X^*$, l'image d'un mot réduit par $T(w)$ est encore réduit. Soit $\rho \in \mathcal{M}_X$ un mot réduit. Écrivons $\rho = w'^*\rho'$, avec $w' \in X \cup X^*$ et $\rho' \in \mathcal{M}_X$. Puisque ρ est un mot réduit, ρ' est réduit.

Si $w \neq w'$, on a $T(w)(\rho) = ww'^*\rho'$. Comme $w'^*\rho'$ est réduit par hypothèse, le seul sous-mot de la forme xx^* ou x^*x (avec $x \in X$) est nécessairement le

sous-mot ww'^*. Or, ce n'est pas possible, puisque $w' \neq w$. Ainsi, $T(w)(\rho)$ est réduit dans ce cas.

Si $w = w'$, on a $T(w)(\rho) = \rho'$, qui est encore un mot réduit.

Montrons maintenant le résultat voulu par récurrence sur la longueur k du mot ω. Si $k = 0$, alors $\omega = \emptyset$, et $T(\omega)(\rho) = \rho$ est réduit. Supposons que le résultat soit vrai pour tout mot de longueur $k \geqslant 0$ et soit ω un mot de longueur $k+1$. On écrit $\omega = w\omega'$, où $w \in X \cup X^*$ et ω' est un mot de longueur k. Par hypothèse de récurrence, $T(\omega')(\rho)$ est réduit. Mais alors, le mot $T(\omega)(\rho) = T(w)\big(T(\omega')(\rho)\big)$ est réduit. Ainsi, la propriété est vraie pour tout mot ω de longueur $k+1$, ce qui achève la récurrence.

(2) On commence par remarquer que $T(w)(\rho) \sim_X w\rho$ pour tout mot ρ.

En effet, c'est clair si ρ ne débute pas par w^*, puisque dans ce cas, on a

$$T(w)(\rho) = w\rho.$$

Si maintenant $\rho = w^*\rho'$, on a

$$w\rho = ww^*\rho' \ \sim_X \ \rho' = T(w)(\rho).$$

Une récurrence immédiate montre alors que l'on a $T(\omega)(\rho) \sim_X \omega\rho$ pour tous mots $\omega, \rho \in \mathcal{M}_X$.

(3) On raisonne encore par récurrence sur la longueur k de ω. C'est vrai si ω est le mot vide. Supposons que le résultat soit vrai pour tout mot de longueur $k \geqslant 0$, et soit ω un mot de longueur $k+1$. On écrit encore $\omega = w\omega'$, où ω' est de longueur k, et $w \in X \cup X^*$. On a $T(\omega)(\emptyset) = T(w)\big(T(\omega')(\emptyset)\big)$. Comme ω est réduit, ω' est réduit et ne débute pas par w^*. Par hypothèse de récurrence, $T(\omega')(\emptyset) = \omega'$, et alors

$$T(\omega)(\emptyset) = T(w)(\omega') = w\omega' = \omega,$$

car ω' ne débute pas par w^*. Cela achève la récurrence.

(4) On commence par démontrer que, pour tout $w \in X \cup X^*$, et pour tout mot réduit ρ, on a

$$\big(T(w) \circ T(w^*)\big)(\rho) = \rho.$$

Soit ρ un mot réduit. Si ρ ne débute pas par w, on a $T(w^*)(\rho) = w^*\rho$. Puisque ce mot débute par w^*, on a alors

$$\big(T(w) \circ T(w^*)\big)(\rho) = T(w)(w^*\rho) = \rho.$$

Si $\rho = w\rho'$, alors $T(w^*)(\rho) = \rho'$. Puisque ρ est réduit, ρ' ne débute pas par le mot w^*, et ainsi

$$\big(T(w) \circ T(w^*)\big)(\rho) = T(w)(\rho') = w\rho' = \rho.$$

Remarquons maintenant que pour tous mots ω_1, ω_2 et tout $w \in X \cup X^*$, on a

$$T(\omega_1 w w^* \omega_2) = T(\omega_1) \circ \big(T(w) \circ T(w^*)\big) \circ T(\omega_2).$$

Par (1), $T(\omega_2)(\emptyset)$ est réduit. Les considérations précédentes montrent alors que

$$\big(T(w)\circ T(w^*)\big)\big(T(\omega_2)(\emptyset)\big) = T(\omega_2)(\emptyset).$$

Finalement, on a

$$T(\omega_1 w w^* \omega_2)(\emptyset) = T(\omega_1)\big(T(\omega_2)(\emptyset)\big) = T(\omega_1\omega_2)(\emptyset).$$

Il s'ensuit que si deux mots ω et ω' sur X sont élémentairement équivalents, alors $T(\omega)(\emptyset) = T(\omega')(\emptyset)$. On en déduit aisément le résultat voulu.

Nous pouvons maintenant achever la démonstration. Soit ω un mot sur X. Par (2), ω est équivalent à $T(\omega)(\emptyset)$, qui est un mot réduit par (1). Si maintenant ω_1 et ω_2 sont deux mots réduits tels que $\omega \sim_X \omega_i$ pour $i = 1, 2$, alors $\omega_1 \sim_X \omega_2$. Par (4), on a $T(\omega_1)(\emptyset) = T(\omega_2)(\emptyset)$. Puisque ω_1 et ω_2 sont réduits, on obtient $\omega_1 = \omega_2$ par le point (3). Ainsi, ω est équivalent à un unique mot réduit, à savoir $T(\omega)(\emptyset)$. Enfin, si ω_1 et ω_2 sont deux mots réduits équivalents, puisque $\omega_1 \sim_X \omega_1$ et $\omega_1 \sim_X \omega_2$, ce qui précède montre que $\omega_1 = \omega_2$. Cela achève la démonstration. □

Nous sommes en mesure maintenant de démontrer le résultat attendu.

4.32. Théorème. *Soit X un ensemble. Alors, (F_X, ι_X) est un groupe libre sur X.*

Démonstration. On sait déjà que F_X est un groupe, engendré par $\iota_X(X)$. Il reste à voir que $\iota_X : X \longrightarrow F_X$ est une application injective, et que la famille $\iota_X(X) = (\overline{x})_{x\in X}$ est libre. Si $x_1, x_2 \in X$ vérifient $\iota_X(x_1) = \iota_X(x_2)$, alors $\overline{x}_1 = \overline{x}_2$. Ainsi, on a $x_1 \sim_X x_2$. Puisque x_1 et x_2 sont réduits, la proposition précédente implique que $x_1 = x_2$, d'où l'injectivité de ι_X.

Soit $k \geqslant 1$, soient $n_1, \ldots, n_k \in \mathbb{Z}\setminus\{0\}$, et soient $x_1, \ldots, x_k \in X$ tels que l'on ait $\overline{x}_i \neq \overline{x}_{i+1}$ pour tout $i \in [\![1, k-1]\!]$. On doit vérifier que $\overline{x}_1^{n_1} \cdots \overline{x}_k^{n_k} \neq \overline{\emptyset}$. L'injectivité de ι_X implique que $x_i \neq x_{i+1}$ pour tout $i \in [\![1, k-1]\!]$.

Posons $w_i = x_i$ si $n_i > 0$ et si $w_i = x_i^*$ si $n_i < 0$. Si l'on note $m_i = |n_i|$, on a alors

$$\overline{x}_1^{n_1} \cdots \overline{x}_k^{n_k} = \overline{w_1^{m_1} \cdots w_k^{m_k}}.$$

Remarquons que $\omega = w_1^{m_1} \cdots w_k^{m_k}$ est réduit. En effet, les sous mots de longueur 2 de ω sont de la forme

$$x_i x_i, \quad x_i^* x_i^*, \quad x_i x_{i+1}, \quad x_i^* x_{i+1}, \quad x_i x_{i+1}^* \quad \text{ou} \quad x_i^* x_{i+1}^*.$$

Aucun de ces sous-mots n'est de la forme xx^* ou x^*x pour $x \in X$, par choix des x_i. Notons que $\omega \neq \emptyset$. Comme ω et $\emptyset$ sont des mots réduits, on en déduit que $\overline{\omega} \neq \overline{\emptyset}$ par la proposition précédente. Ainsi, $\iota_X(X)$ est libre. Cela achève la démonstration. □

5. Exercices

E1. On considère le sous-groupe de $\mathbb{Z}^2$

$$H = \{(x, y) \in \mathbb{Z}^2 \mid x + y \in 2\mathbb{Z},\ x - y \in 3\mathbb{Z}\}.$$

Lister les éléments de $\mathbb{Z}^2/H$ et donner la table de la loi de groupe.

E2. On considère le sous-groupe $H = 4\mathbb{Z} \times 6\mathbb{Z}$ de $\mathbb{Z}^2$.

a. Montrer que $\mathbb{Z}^2/H$ est fini et donner son nombre d'éléments.

b. Déterminer l'ordre de $\overline{(6, -4)}$ dans $\mathbb{Z}^2/H$.

c. Le groupe $\mathbb{Z}^2/H$ est-il cyclique ?

E3. On considère le groupe quotient $\mathbb{R}/\mathbb{Z}$.

a. Peut-on définir une loi produit sur $\mathbb{R}/\mathbb{Z}$ par $\overline{x} \cdot \overline{y} = \overline{xy}$?

b. Montrer que $\mathbb{R}/\mathbb{Z}$ est infini.

Indication. Si $\mathbb{R}/\mathbb{Z}$ était fini, $\mathbb{R}$ serait la réunion disjointe d'un nombre fini de classes à gauche. Que dire alors de la cardinalité de $\mathbb{R}$?

c. Montrer que $\mathbb{R}/\mathbb{Z}$ n'a qu'un nombre fini d'éléments d'ordre 2 et 3, et en donner la liste. Quels sont les éléments d'ordre fini ? Donner l'exemple d'un élément d'ordre infini. Exhiber un élément d'ordre $n \geqslant 1$ (n fixé).

d. Quel est le sous-groupe de $\mathbb{R}/\mathbb{Z}$ engendré par $\overline{\left(\frac{7}{12}\right)}$?

e. Montrer que $\mathbb{Q}/\mathbb{Z}$ est un sous-groupe infini de $\mathbb{R}/\mathbb{Z}$, qui n'a que des éléments d'ordre fini.

f. Le groupe $\mathbb{R}/\mathbb{Z}$ est-il monogène ?

E4. Soit $f : G \longrightarrow G'$ un morphisme de groupes.

a. Soit H' un sous-groupe distingué de G'. Montrer que $f^{-1}(H')$ est un sous-groupe distingué de G.

b. On suppose que f est surjective. Soit H un sous-groupe distingué de G. Montrer que $f(H)$ est un sous-groupe distingué de G'. Et si f n'est pas supposée surjective ?

E5. Soit G un groupe, soit H un sous-groupe distingué de G et soit K un sous-groupe distingué de H. Le sous-groupe K est-il nécessairement distingué dans G ?

On pourra considérer $G = D_4$.

E6.

a. Montrer que le groupe $\mathbb{C}/\mathbb{R}$ est isomorphe à $\mathbb{R}$.

b. Montrer que $\mathbb{R}^\times/\mathbb{R}^{\times 2} \simeq \{\pm 1\}$, où $\mathbb{R}^{\times 2} = \{x^2 \mid x \in \mathbb{R}^\times\}$.

E7. Soit $f : G \longrightarrow G'$ un morphisme de groupes, et soit H un sous-groupe distingué de G contenu dans $\mathrm{Ker}(f)$. Montrer que $\overline{f} : G/H \longrightarrow G'$ est injectif si, et seulement si, $H = \mathrm{Ker}(f)$.

E8. Décrire tous les morphismes de groupes $\mathbb{Z}/m\mathbb{Z} \longrightarrow \mathbb{Z}/n\mathbb{Z}$. Quelle est la structure du groupe $\mathrm{Hom}(\mathbb{Z}/m\mathbb{Z}, \mathbb{Z}/n\mathbb{Z})$?

E9. Soit G un groupe. On rappelle que le centre de G est le sous-groupe

$$Z(G) = \{z \in G \mid zg = gz, \text{ pour tout } g \in G\}.$$

a. Montrer que $Z(G)$ est distingué dans G, abélien, et que l'on a un isomorphisme de groupes

$$G/Z(G) \simeq \mathrm{Int}(G),$$

où $\mathrm{Int}(G)$ est le sous-groupe de $\mathrm{Aut}(G)$ constitué des automorphismes intérieurs.

b. On suppose que $G/Z(G)$ est un groupe monogène. Montrer que G est abélien, et donc que $G = Z(G)$.

E10. Soit G un groupe fini. On pose

$$R(G) = \{(x, y) \in G \times G \mid xy = yx\},$$

ainsi que $r(G) = \dfrac{|R(G)|}{|G|^2}$, appelé *ratio de commutativité de* G.

On se propose dans cet exercice d'étudier quelques propriétés de cet entier dépendant de G.

a. Soit E un ensemble fini non vide, muni d'une relation d'équivalence $\sim$. Montrer que l'on a

$$\sum_{x \in E} \frac{1}{|\overline{x}|} = |E/\!\sim|.$$

b. Soit G un groupe fini agissant sur un ensemble non vide E. On note Ω l'ensemble des orbites de E sous l'action de G. Déduire de la question précédente que l'on a

$$\frac{1}{|G|^2} \sum_{x \in E} |\mathrm{Stab}_G(x)| = \frac{|\Omega|}{|G|}.$$

c. Montrer que $0 < r(G) \leqslant 1$. À quelles conditions nécessaires et suffisantes sur G a-t-on $r(G) = 1$?

d. Pour tout $x \in G$, on pose $C_G(x) = \{y \in G \mid xy = yx\}$.

Vérifier que $|R(G)| = \sum_{x \in G} |C_G(x)|$.

En utilisant les questions précédentes, en déduire que

$$r(G) = \frac{c(G)}{|G|},$$

où $c(G)$ est le nombre de classes de conjugaison de G.

e. Calculer $r(\mathfrak{S}_7)$, ainsi que $r(D_4)$.

f. Soient G_1, G_2 deux groupes finis. Montrer que $r(G_1 \times G_2) = r(G_1)r(G_2)$. Que vaut $\inf_{G \text{ fini}} r(G)$?

g. En utilisant l'exercice précédent, montrer qu'un groupe G non abélien d'ordre p^3, où p est premier, vérifie $[G : Z(G)] = p^2$.

h. En déduire également que si G est un groupe fini non abélien, alors on a l'inégalité $\dfrac{|G|}{|Z(G)|} \geqslant 4$.

i. Si G un groupe fini, justifier qu'il y a $|Z(G)|$ classes de conjugaison à un élément, puis que

$$c(G) \leqslant |Z(G)| + \frac{|G| - |Z(G)|}{2}.$$

j. En déduire que si G est non abélien, on a

$$r(G) \leqslant \frac{1}{2}\left(1 + \frac{|Z(G)|}{|G|}\right) \leqslant \frac{5}{8}.$$

k. Soit G un groupe fini non abélien. Montrer que $r(G) = \dfrac{5}{8}$ si, et seulement si, $[G : Z(G)] = 4$.

Donner un exemple concret de tel groupe, et calculer explicitement son centre.

l. Donner un exemple de famille infinie de groupes finis $(G_n)_{n \geqslant 0}$ telle que $r(G_n) = \dfrac{5}{8}$, pour tout $n \geqslant 0$.

E11. On considère le groupe Q_8, défini dans l'exercice VI-11.

a. Vérifier que tous les sous-groupes de Q_8 sont distingués dans Q_8. Le groupe Q_8 est-il abélien ?

b. Identifier chaque quotient Q_8/H à un groupe connu.

E12. Soient G_1, G_2 deux groupes, et soit H_i un sous-groupe distingué dans G_i, $i = 1, 2$. Montrer que $H_1 \times H_2$ est distingué dans $G_1 \times G_2$, et que l'on a un isomorphisme

$$(G_1 \times G_2)/(H_1 \times H_2) \simeq G_1/H_1 \times G_2/H_2.$$

E13. Soit G un groupe fini, et soient H_1, H_2 deux sous-groupes distingués dans G.

a. On suppose que $H_1 \simeq H_2$. A-t-on $G/H_1 \simeq G/H_2$?

b. On suppose que $G/H_1 \simeq G/H_2$. A-t-on $H_1 \simeq H_2$?

Indication. Pour les deux questions, on pourra considérer un groupe abélien d'ordre 8 bien choisi.

E14. Soit G un groupe (noté multiplicativement) et soit B un groupe **abélien** noté additivement.

On suppose que G agit sur B. On dit que l'action de G sur B est une action *par automorphismes* si

$$g{\cdot}(b + b') = g{\cdot}b + g{\cdot}b', \quad \text{pour tout } g \in G \text{ et tous } b, b' \in B.$$

Si B, C sont deux groupes abéliens sur lesquels G agit par automorphismes, on dit qu'un morphisme de groupes $f : B \longrightarrow C$ est G-**équivariant** si l'on a

$$f(g{\cdot}b) = g{\cdot}f(b), \quad \text{pour tout } g \in G \text{ et tous } b \in B.$$

Dans toute la suite, G agit par automorphismes sur les groupes abéliens B et C.

a. Soit $f : B \longrightarrow C$ un morphisme de groupes G-équivariants. Montrer que $f(B^G) \subset C^G$.

On note $f_* : B^G \longrightarrow C^G$ le morphisme de groupes induit par f par double restriction (on a donc $f_*(b) = f(b)$ pour tout $b \in B^G$).

b. Montrer que l'action de G sur B induit une action de G sur $\mathrm{Ker}(f)$, et que l'on a $\mathrm{Ker}(f_*) = \big(\mathrm{Ker}(f)\big)^G$.

c. Montrer que si f est injective, resp. un isomorphisme , alors f_* est injective, resp. un isomorphisme.

d. On fait agir $G = \{\pm 1\}$ sur $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$ par

$$(-1)^m{\cdot}(\overline{x}, \overline{y}) = (\overline{x} + \overline{m}\ \overline{y}, \overline{y}), \quad \text{pour tous } \overline{x}, \overline{y}, \overline{m} \in \mathbb{Z}/2\mathbb{Z}.$$

On fait aussi agir G sur $\mathbb{Z}/2\mathbb{Z}$ trivialement.

Montrer que ces deux actions sont des actions par automorphismes, et que le morphisme de groupes

$$\begin{array}{rcl} f : \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} & \longrightarrow & \mathbb{Z}/2\mathbb{Z} \\ (\overline{x}, \overline{y}) & \longmapsto & \overline{y} \end{array}$$

est G-équivariant et surjectif. Le morphisme f_* est-il surjectif ?

Le but de ce qui suit est de trouver une condition nécessaire et suffisante pour que f_* soit surjective. On commence par introduire un groupe quotient qui nous servira par la suite.

Soit A un groupe abélien (noté additivement), sur lequel G agit par automorphismes.
On note $Z^1(G, A)$ l'ensemble des applications $\alpha : G \longrightarrow A$ telles que

$$\alpha(1) = 0 \text{ et } \alpha(gg') = \alpha(g) + g{\cdot}\alpha(g'), \quad \text{pour tous } g, g' \in G.$$

On note $B^1(G, A)$ l'ensemble des applications $\beta : G \longrightarrow A$ telles qu'il existe $a \in A$ vérifiant

$$\beta(g) = a - g{\cdot}a, \quad \text{pour tout } g \in G.$$

e. Montrer que $Z^1(G, A)$ est un groupe abélien, et que $B^1(G, A)$ est un sous-groupe de $Z^1(B, A)$.

On note $H^1(G, A)$ le groupe quotient $Z^1(B, A)/B^1(G, A)$.
Si $\alpha \in Z^1(G, A)$, on note $[\alpha]$ sa classe dans $H^1(G, A)$.

f. On suppose que A est un groupe fini.
Montrer que pour tout $[\alpha] \in H^1(G, A)$, on a $o([\alpha]) \mid |A|$.

g. On suppose que G est un groupe fini.
Montrer que pour tout $[\alpha] \in H^1(G, A)$, on a $o([\alpha]) \mid |G|$.

Indication. Sommer sur g' la relation $\alpha(gg') = \alpha(g) + g{\cdot}\alpha(g')$.

h. On suppose que G et A sont des groupes finis d'ordres premiers entre eux. Montrer que $H^1(G, A)$ est le groupe trivial.

On revient maintenant au problème initial.

Soit $f : B \longrightarrow C$ un morphisme de groupes abéliens G-équivariant et **surjectif**. On pose $A = \mathrm{Ker}(f)$.

Soit $c \in C^G$. On se fixe un élément $b \in B$ tel que $f(b) = c$ (ce qui est toujours possible car f est surjectif).

i. Montrer que pour tout $g \in G$, on a $b - g{\cdot}b \in A$, et que l'application

$$\begin{array}{rcl} \alpha_c : G & \longrightarrow & A \\ g & \longmapsto & b - g{\cdot}b \end{array}$$

appartient à $Z^1(G, A)$.

j. Soit $b' \in B$ un autre élément tel que $f(b') = c$. On note α'_c l'application

$$\begin{aligned} \alpha'_c : G &\longrightarrow A \\ g &\longmapsto b' - g{\cdot}b'. \end{aligned}$$

Montrer que $\alpha_c - \alpha'_c \in B^1(G, A)$.

k. En déduire que la classe $[\alpha_c] \in H^1(G, A)$ ne dépend pas du choix de b, et que l'application

$$\begin{aligned} \delta : C^G &\longrightarrow H^1(G, A) \\ c &\longmapsto [\alpha_c] \end{aligned}$$

est un morphisme de groupes bien défini.

l. Montrer que $\mathrm{Im}(f_*) = \mathrm{Ker}(\delta)$ (procéder par double inclusion). En déduire que f_* est surjective si, et seulement si, δ est le morphisme nul.

m. On reprend l'exemple de la question **d**. Calculer $\delta\big((\overline{0}, \overline{1})\big)$ et montrer que c'est un élément non nul de $H^1(G, A)$.

n. On suppose que G et B sont deux groupes finis d'ordre premiers entre eux. Montrer que f_* est surjective.

o. On suppose B et C sont deux groupes finis d'ordre premiers entre eux. Montrer que f_* est surjective.

E15. Soit G un groupe. Un sous-groupe H de G est dit *maximal* si $H \neq G$, et si pour tout sous-groupe H' de G, on a

$$H \subset H' \implies H = H' \text{ ou } H = G.$$

a. On suppose que $[G : H]$ est fini, et est un nombre premier. Montrer que H est un sous-groupe maximal de G.

b. Montrer que tout groupe fini non réduit au neutre possède au moins un sous-groupe maximal. En considérant $G = \mathfrak{A}_4$, montrer que l'indice d'un sous-groupe maximal n'est pas nécessairement premier.

c. On suppose que G possède un sous-groupe maximal H distingué dans G. Montrer que H est d'indice fini dans G, et que $[G : H]$ est un nombre premier.

Indication. Montrer que dans ce cas, les seuls sous-groupes de G/H sont $\{\overline{1}_G\}$ et G/H.

d. Montrer que le seul sous-groupe de $\mathbb{Q}$ d'indice fini est $\mathbb{Q}$. En déduire que $\mathbb{Q}$ ne possède pas de sous-groupes maximaux.

Indication. Si H est un sous-groupe de $\mathbb{Q}$ d'indice fini, montrer qu'il existe $n \geqslant 1$ tel que $n\mathbb{Q} \subset H$.

E16. Soit G un p-groupe, et soit $H \neq G$ un sous-groupe.

a. Montrer l'existence d'un sous-groupe de G d'indice p contenant H.

Indication. Procéder par récurrence sur $|G|$.

Si H contient un élément $a \in Z(G)$ non trivial, appliquer l'hypothèse de récurrence à $G/\langle a \rangle$.

Si $H \cap Z(G) = \{1_G\}$, choisir un élément $a \in Z(G)$ d'ordre p (après avoir justifié son existence) et considérer $H' = \langle H, a \rangle$. Si H n'est pas d'indice p, montrer que $H' \neq G$ et contient un élément central non trivial.

b. En déduire l'existence d'une chaîne de sous-groupes

$$H = H_0 \subset H_1 \subset \cdots \subset H_{r-1} \subset H_r = G$$

vérifiant $[H_i : H_{i-1}] = p$ pour tout $i \in [\![1, r]\!]$.

E17. Soit G un groupe. On note $[G, G]$ le sous-groupe de G engendré par les éléments

$$[g_1, g_2] = g_1 g_2 g_1^{-1} g_2^{-1}, \quad g_1, g_2 \in G.$$

a. Montrer que $[G, G]$ est distingué dans G, et que $G/[G, G]$ est abélien.

b. Si H est un sous-groupe distingué dans G tel que G/H est abélien, montrer que $[G, G] \subset H$. Inversement, montrer que si H est un sous-groupe contenant $[G, G]$, alors H est distingué dans G et G/H est abélien.

c. Montrer que tout morphisme de groupes $G \longrightarrow A$, avec A abélien, se factorise en un morphisme de groupes $G/[G, G] \longrightarrow A$.

d. Soit G^2 le sous-groupe de G engendré par les éléments g^2, $g \in G$. Montrer que $[G, G] \subset G^2$.

e. Soit p un nombre premier, et soit G^* le sous-groupe engendré par $[G, G]$ et les éléments g^p, $g \in G$.

(i) Montrer que G^* est distingué dans G, que G/G^* est abélien, et tué par p, i.e. pour tout $\overline{g} \in G/G^*$, on a $\overline{g}^p = \overline{1}$. En déduire que si G est fini, on a

$$G/G^* \simeq (\mathbb{Z}/p\mathbb{Z})^r,$$

pour un certain $r \geqslant 0$.

Indication. Utiliser l'exercice VI-43.

(ii) Pour tout sous-groupe H distingué dans G tel que G/H soit abélien, tué par p, montrer que $G^* \subset H$.

f. On suppose dans cette question que G est un p-groupe. On rappelle[8] que tout sous-groupe $H \neq G$ est contenu dans un sous-groupe H' d'indice p et que H' est nécessairement distingué dans G, par le théorème de Frobenius (théorème VII-2.4).

Soient $\overline{x}_1, \ldots, \overline{x}_r$ des générateurs de G/G^*. Montrer que

$$G = \langle x_1, \ldots, x_r \rangle.$$

On pourra introduire le sous-groupe engendré par les x_i.

g. On suppose que G est non abélien d'ordre p^3. Montrer que

$$G/G^* \simeq (\mathbb{Z}/p\mathbb{Z})^2 \quad \text{et} \quad [G,G] = G^* \simeq \mathbb{Z}/p\mathbb{Z}.$$

E18. On suppose connus les définitions et résultats des exercices 15 et 17.

a. Soit G un groupe, et soit H un sous-groupe distingué de G. Montrer que $\pi : G \longrightarrow G/H$ induit une bijection entre l'ensemble des sous-groupes maximaux de G contenant H et l'ensemble des sous-groupes maximaux de G/H.

b. Montrer qu'un isomorphisme $f : G_1 \xrightarrow{\sim} G_2$ induit une bijection de l'ensemble des sous-groupes maximaux de G_1 sur l'ensemble des sous-groupes maximaux de G_2.

Si G est un groupe fini, on note $\Phi(G)$ l'intersection des sous-groupes maximaux de G.

c. Montrer qu'un isomorphisme $f : G_1 \xrightarrow{\sim} G_2$ induit par double restriction un isomorphisme $\Phi(G_1) \simeq \Phi(G_2)$.

d. En déduire que $\Phi(G)$ est un sous-groupe distingué de G.

e. Soit $G = (\mathbb{Z}/p\mathbb{Z})^n$, où p est premier et $n \geqslant 1$.
En utilisant la question **a** de l'exercice 15, montrer que $\Phi(G) = \{0_G\}$.

f. Soit G un groupe, et soit H un sous-groupe distingué de G.
On suppose que $G/H \simeq (\mathbb{Z}/p\mathbb{Z})^n$, où p est premier et $n \geqslant 1$. Montrer l'égalité

$$H = \bigcap_{\substack{H \subset M \\ M \text{ maximal}}} M.$$

On s'intéresse maintenant aux sous-groupes maximaux des p-groupes finis. Jusqu'à la fin de l'exercice, G est un groupe d'ordre p^n, où p est premier et $n \geqslant 1$.

8. Cf. exercice 16.

g. Montrer que les sous-groupes maximaux de G sont exactement les sous-groupes d'indice p. On pourra utiliser l'exercice 16.

Quel est le théorème qui permet de dire qu'un tel sous-groupe maximal est distingué dans G ?

h. Soient $M_1, \ldots, M_r$ les sous-groupes maximaux de G. En considérant le morphisme

$$\begin{aligned} f : G &\longrightarrow G/M_1 \times \cdots \times G/M_r \\ g &\longmapsto (gM_1, \ldots, gM_r), \end{aligned}$$

montrer soigneusement que le groupe quotient $G/\Phi(G)$ est abélien, et que pour tout $g \in G$, on a $\overline{g}^p = \overline{1} \in G/\Phi(G)$.

i. Déduire des questions précédentes que $\Phi(G) = G^*$.

Les exercices suivants nécessitent d'avoir étudié les sections 3 et 4.

E19. Soit G un p-groupe tel que $G/[G,G] \simeq \mathbb{Z}/p\mathbb{Z}$. Montrer que $G \simeq \mathbb{Z}/p\mathbb{Z}$.

Indication. Si $p^2 \mid |G|$, considérer un sous-groupe H distingué dans G tel que $[G : H] = p^2$, et montrer que $[G,G] \subset H$. On pourra utiliser les résultats de l'exercice IX-17.

E20. Démontrer le théorème fondamental de l'arithmétique en utilisant le théorème de Jordan-Hölder pour un groupe abélien bien choisi.

E21. Donner une suite de Jordan Hölder pour les groupes $\mathfrak{S}_3$, $\mathfrak{S}_4$ et D_n, et identifier les quotients successifs.

E22. Soient G_1, G_2 deux groupes. Montrer que $G_1 \times G_2$ est résoluble si, et seulement si, G_1 et G_2 sont résolubles.

E23. On suppose connus les résultats de l'exercice 17, dont on reprend les notations. Soit G un groupe. Pour tout $n \geqslant 0$, on pose

$$G^{(0)} = G \quad \text{et} \quad G^{(n+1)} = [G^{(n)}, G^{(n)}].$$

a. Pour tout sous-groupe H distingué dans G, montrer que $[H,H]$ est distingué dans G, et que $H/[H,H]$ est abélien.

b. Montrer l'équivalence des propriétés suivantes :

(1) il existe $n \geqslant 0$ tel que $G^{(n)} = \{1_G\}$;

(2) il existe une suite de composition $(G_0, \ldots, G_n)$ de G telle que pour tout $i \in [\![0, n-1]\!]$, G_i est distingué dans G et G_i/G_{i+1} est abélien ;

(3) le groupe G est résoluble.

Indication. Pour (1) $\Longrightarrow$ (2), on pourra prendre $G_i = G^{(i)}$. Pour démontrer (3) $\Longrightarrow$ (1), si $(G_0, \ldots, G_r)$ est une suite de composition dont les quotients successifs sont abéliens, on pourra d'abord montrer l'égalité $G_{i+1} = G_i^{(1)}$ pour tout $i \in [\![0, r-1]\!]$.

E24. Soit G un groupe fini résoluble non réduit au neutre.

a. Montrer qu'il existe un sous-groupe H distingué dans G, minimal pour l'inclusion. Autrement dit, H est distingué dans G, $H \neq \{1_G\}$, et le seul sous-groupe strict de H qui soit distingué dans G est $\{1_G\}$. Un tel sous-groupe H est-il unique à isomorphisme près ?

b. Montrer que $[H, H] \neq H$, et en déduire que H est abélien.

c. Montrer que pour tout p premier, $H(p) = \{x \in H \mid x^p = 1_G\}$ est un sous-groupe distingué de G.

d. En déduire qu'il existe un nombre premier p et un entier $r \geqslant 1$ tel que l'on ait $H \simeq (\mathbb{Z}/p\mathbb{Z})^r$.

Indication. Montrer qu'il existe un premier p tel que $H(p) \neq \{1_G\}$, et utiliser l'exercice VI-43.

E25. Soit G un groupe, et soient H et K deux sous-groupes distingués dans G.

a. Montrer que si H et K sont résolubles, HK est résoluble.

b. Montrer que si G/H et G/K sont résolubles, $G/(H \cap K)$ est résoluble.

E26. Soit G un groupe fini.

a. Montrer qu'il existe un unique sous-groupe distingué K de G vérifiant les propriétés suivantes :

(1) G/K est résoluble ;

(2) tout sous-groupe N distingué dans G tel que G/N est résoluble contient K.

b. Montrer que pour tout $f \in \mathrm{Aut}(G)$, on a $f(K) = K$.

c. Montrer que $K = [K, K]$. En déduire que, soit $K = \{1_G\}$, soit K n'est pas résoluble.

E27. Trouver une présentation par générateurs et relations des groupes suivants :

$$\mathbb{Z}^2, \quad \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}, \quad Q_8.$$

E28. Soit $n \geqslant 3$. Soit $Q_{2^n} = \langle a, b \mid a^{2^{n-1}} = 1,\ a^{2^{n-2}} = b^2,\ bab^{-1} = a^{-1} \rangle$.

a. Montrer que tout élément de Q_{2^n} s'écrit sous la forme

$$a^i b^j, \quad i \in [\![0, 2^{n-1} - 1]\!],\ j = 0, 1.$$

b. Montrer que Q_{2^n} possède un quotient isomorphe à $D_{2^{n-2}}$ (avec la convention $D_2 = \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$). En déduire que Q_{2^n} n'est pas abélien.

c. En déduire que $|Q_{2^n}| = 2^n$.

E29. Soit $n \geqslant 2$, et soit

$$G_n = \langle t_1, \ldots, t_{n-1} \mid t_i^2 = 1,\ (t_i t_{i+1})^3 = 1,\ t_i t_j = t_j t_i \text{ si } |j - i| \geqslant 2 \rangle.$$

Le but de cet exercice est de montrer que $G_n \simeq \mathfrak{S}_n$.

a. Montrer qu'il existe un morphisme surjectif $\varphi : G_n \longrightarrow \mathfrak{S}_n$. En déduire que $|G_n| \geqslant n!$

Soit $H = \langle t_2, \ldots, t_{n-1} \rangle$. Pour tout $i \in [\![0, n-1]\!]$, on pose

$$\sigma_i = t_i t_{i-1} \cdots t_2 t_1 \quad \text{et} \quad H_i = \sigma_i H.$$

b. On fixe $i \in [\![0, n-1]\!]$.

(i) Si $i \geqslant 1$, montrer que $t_i H_i = H_{i-1}$.

(ii) Si $i \leqslant n-2$, montrer que $t_{i+1} H_i = H_{i+1}$.

(iii) Montrer que pour tout $j \geqslant i+2$, on a $t_j H_i = H_i$.

(iv) Soit $j \leqslant i-1$. Montrer que tout élément du groupe G_n s'écrit sous la forme $\sigma t_{j+1} t_j \sigma_j$, où σ commute avec t_j et t_{j+1}. En utilisant la relation $(t_j t_{j+1})^3 = 1$, en déduire que $t_j H_i = H_i$.

c. En déduire que pour tout $g \in G_n$, il existe un entier $i \in [\![0, n-1]\!]$ tel que $gH = \sigma_i H$, puis que $[G_n : H] \leqslant n$.

d. Montrer par récurrence que $|G_n| \leqslant n!$, et conclure.

E30. On se propose de retrouver l'existence d'un automorphisme non intérieur de $\mathfrak{S}_6$ (cf. exercice VIII-32).

Soient $\tau_1, \ldots, \tau_5 \in \mathfrak{S}_6$ les éléments de $\mathfrak{S}_6$ suivants :

$$\tau_1 = (1\ 3)(2\ 4)(5\ 6), \quad \tau_2 = (1\ 2)(3\ 6)(4\ 5), \quad \tau_3 = (1\ 4)(2\ 3)(5\ 6),$$

$$\tau_4 = (1\ 3)(2\ 6)(4\ 5), \quad \tau_5 = (1\ 2)(3\ 4)(5\ 6).$$

a. À l'aide de l'exercice précédent, montrer qu'il existe un unique morphisme $\varphi : \mathfrak{S}_6 \longrightarrow \mathfrak{S}_6$ qui envoie $(i\ i+1)$ sur τ_i, pour tout $i \in [\![1, 5]\!]$.

b. Vérifier que $\mathrm{Im}(\varphi)$ n'est pas abélien. En déduire que $\mathrm{Ker}(\varphi)$ est différent de $\mathfrak{A}_6$ et de $\mathfrak{S}_6$, puis que φ est un automorphisme.

c. Justifier que φ est un automorphisme de $\mathfrak{S}_6$ non intérieur (on pourra utiliser les résultats de l'exercice VIII-31).

Chapitre X

Produits directs et semi-directs

1. Préliminaires

Si G un groupe, et si $H_1, \ldots, H_r$ sont des sous-groupes de G, on pose

$$H_1 \cdots H_r = \{h_1 \cdots h_r \mid h_i \in H_i\}.$$

Ce n'est pas un groupe en général.

On note $\langle H_1, \ldots, H_r \rangle$ le sous-groupe de G engendré par $H_1, \ldots, H_r$ (c'est-à-dire par le sous-ensemble $H_1 \cup \cdots \cup H_r$). On a clairement

$$H_1 \cdots H_r \subset \langle H_1, \ldots, H_r \rangle,$$

mais l'inclusion est stricte en général.

Le lemme suivant généralise la première partie du théorème IX-2.5.

1.1. Lemme. *Soit G un groupe, et soient $H_1, \ldots, H_r$ des sous-groupes de G, avec $r \geqslant 2$. On suppose que $H_1, \ldots, H_{r-1}$ sont distingués dans G. Alors, l'ensemble $H_1 \cdots H_r$ est un sous-groupe de G, et l'on a*

$$\langle H_1, \ldots, H_r \rangle = H_1 \cdots H_r.$$

Démonstration. Pour vérifier que $H_1 \cdots H_r$ est bien un sous-groupe de G, on procède par récurrence sur r. Le cas $r = 2$ n'est rien d'autre que la première partie du théorème IX-2.5. Supposons maintenant que le résultat soit vrai pour $r \geqslant 2$, et considérons $r+1$ sous-groupes $H_1, \ldots, H_{r+1}$ dont les r premiers sont distingués dans G. L'hypothèse de récurrence montre que $H_2 \cdots H_{r+1}$ est un sous-groupe de G. De plus, H_1 étant un sous-groupe

distingué de G, le cas $r = 2$ nous donne que $H_1(H_2 \cdots H_{r+1}) = H_1 \cdots H_{r+1}$ est un sous-groupe de G. Par conséquent, la propriété désirée est vraie au rang $r + 1$. Cela achève la récurrence.

Mais alors, l'ensemble $H_1 \cdots H_r$ est un sous-groupe de G, qui contient clairement $H_1, \ldots, H_r$. Ainsi, $\langle H_1, \ldots, H_r \rangle \subset H_1 \cdots H_r$. L'autre inclusion étant évidente, cela achève la démonstration. □

On finit par un lemme qui nous sera utile par la suite.

1.2. Lemme. *Soit G un groupe, et soient H, K deux sous-groupes finis de G. Alors, HK est un ensemble fini, et l'on a*

$$|HK| = \frac{|H| \times |K|}{|H \cap K|}.$$

Démonstration. Considérons l'application

$$\begin{aligned} f \colon H \times K &\longrightarrow HK \\ (h, k) &\longmapsto hk. \end{aligned}$$

Puisque f est surjective, on a

$$|HK| \leqslant |H \times K| < +\infty,$$

et donc HK est fini. De plus, l'ensemble $H \times K$ est la réunion disjointe des ensembles $f^{-1}(\{hk\})$, lorsque hk parcourt HK. Or, pour tout $h \in H$ et tout $k \in K$, on a

$$f^{-1}(\{hk\}) = \{(hg, g^{-1}k) \mid g \in H \cap K\}.$$

En effet, l'inclusion « $\supset$ » est évidente. Si maintenant $(h', k') \in H \times K$ vérifie $h'k' = hk$, on a $h^{-1}h' = kk'^{-1}$. Mais alors, $g = h^{-1}h' \in H \cap K$, et l'on a

$$hg = h(h^{-1}h') = h' \quad \text{et} \quad g^{-1}k = (kk'^{-1})^{-1}k = k'.$$

De plus, pour tous $g_1, g_2 \in H \cap K$, on a

$$(hg_1, g_1^{-1}k) = (hg_2, g_2^{-1}k) \quad \Longrightarrow \quad g_1 = g_2,$$

comme on peut le voir par simplification dans G. Ainsi, pour tout $hk \in HK$, on a $|f^{-1}(hk)| = |H \cap K|$. On obtient donc

$$|H \times K| = |HK| \times |H \cap K|,$$

d'où la conclusion. □

2. Produits directs

Le but de ce paragraphe est de définir la notion de produit direct interne et externe de groupes, qui peut se comparer à la somme directe interne et externe d'espaces vectoriels. On commence par introduire le produit direct externe.

2.1. Définition. Soient $G_1, \ldots, G_r$ des groupes.

Le groupe $G_1 \times \cdots \times G_r$ (muni de la loi de groupe évidente) est appelé le *produit direct externe* des groupes $G_1, \ldots, G_r$.

2.2. Lemme. *Soit $G = G_1 \times \cdots \times G_r, r \geqslant 2$. Pour $i \in [\![1, r]\!]$, on pose*

$$H_i = \{(1_{G_1}, \ldots, 1_{G_{i-1}}, g_i, 1_{G_{i+1}}, \ldots, 1_{G_r}) \mid g_i \in G_i\}.$$

Alors, on a les propriétés suivantes :

(1) *les ensembles $H_1, \ldots, H_r$ sont des sous-groupes distingués de G ;*

(2) *on a $h_i h_j = h_j h_i$ pour tout $h_i \in H_i$, et tout $h_j \in H_j$, dès que $i \neq j$;*

(3) *tout élément $g \in G$ s'écrit de manière unique sous la forme*

$$g = h_1 \cdots h_r, \quad h_i \in H_i.$$

Démonstration. Exercice laissé au lecteur. □

Ce lemme motive la définition suivante.

2.3. Définition. Soit G un groupe, et soient r sous-groupes $H_1, \ldots, H_r$ de G (avec $r \geqslant 2$). On dit que G est le *produit direct interne* de $H_1, \ldots, H_r$ si les deux conditions suivantes sont réalisées :

(1) on a $h_i h_j = h_j h_i$ pour tout $h_i \in H_i$, et tout $h_j \in H_j$, dès que $i \neq j$;

(2) tout élément $g \in G$ s'écrit de manière unique sous la forme

$$g = h_1 \cdots h_r, \quad h_i \in H_i.$$

On le note $G = H_1 \odot \cdots \odot H_r$.

2.4. Remarque. Soit G un groupe, et soient $H_1, \ldots, H_r$ des sous-groupes de G. Alors, $G = H_1 \odot \cdots \odot H_r$ si, et seulement si, l'application

$$\begin{aligned} H_1 \times \cdots \times H_r &\longrightarrow G \\ (h_1, \ldots, h_r) &\longmapsto h_1 \cdots h_r. \end{aligned}$$

est un isomorphisme. Cela est laissé en exercice au lecteur.

Nous allons maintenant donner des conditions nécessaires et suffisantes pour qu'un groupe G soit un produit direct interne de sous-groupes.

2.5. Proposition. *Soit G un groupe, et soient $H_1, \ldots, H_r$ r sous-groupes de G ($r \geqslant 2$). Alors, $G = H_1 \odot \cdots \odot H_r$ si, et seulement si, les conditions suivantes sont satisfaites :*

(1) *on a $h_i h_j = h_j h_i$ pour tout $h_i \in H_i$, et tout $h_j \in H_j$, dès que $i \neq j$;*

(2) *on a $G = \langle H_1, \ldots, H_r \rangle$;*

(3) *pour tout $i \in [\![1, r-1]\!]$, on a $H_1 \cdots H_i \cap H_{i+1} = \{1_G\}$.*

Démonstration. Supposons que $G = H_1 \odot \cdots \odot H_r$. Alors, (1) est vérifiée par définition. De plus, on a $G = H_1 \cdots H_r \subset \langle H_1, \ldots, H_r \rangle$, et l'on a donc montré (2), puisque l'autre inclusion est triviale. Montrons enfin que (3) est réalisée. Soit $i \in [\![1, r-1]\!]$, et soit $g \in H_1 \cdots H_i \cap H_{i+1}$. Alors, il existe des éléments $h_1 \in H_1, \ldots, h_i \in H_i$ tels que

$$g = h_1 \cdots h_i.$$

Comme on a aussi $g \in H_{i+1}$, l'unicité de la décomposition d'un élément entraîne que tous les h_i sont égaux à 1_G, et donc que $g = 1_G$.

Inversement, supposons les conditions (1) – (3) vérifiées. Il reste à montrer que tout élément $g \in G$ se décompose de manière unique sous la forme

$$g = h_1 \cdots h_r, \quad h_i \in H_i.$$

Remarquons tout d'abord que $H_1, \ldots, H_r$ sont distingués dans G. En effet, soit $h_i \in H_i$ pour tout $i \in [\![1, r]\!]$. Comme on a $G = \langle H_1, \ldots, H_r \rangle$, il suffit de vérifier que l'on a $h_j h_i h_j^{-1}$ pour tout $h_j \in H_j$ et tout entier $j \in [\![1, r-1]\!]$. Si $j = i$, cela provient du fait que H_i est un sous-groupe de G. Si $j \neq i$, cela provient de l'égalité $h_j h_i h_j^{-1} = h_i$.

Le lemme 1.1 implique alors que $G = H_1 \cdots H_r$, c'est-à-dire l'existence de la décomposition recherchée. Pour montrer l'unicité, supposons que l'on ait une égalité de la forme

$$h_1 \cdots h_r = h'_1 \cdots h'_r, \quad \text{avec } h_i,\ h'_i \in H_i.$$

En tenant compte de la condition (1), cela se récrit

$$h_r (h'_r)^{-1} = h_1^{-1} h'_1 \cdots h_{r-1}^{-1} h'_{r-1}.$$

La condition (3) appliquée à $i = r-1$ montre alors que $h_r (h'_r)^{-1} = 1_G$ et que $h_1^{-1} h'_1 \cdots h_{r-1}^{-1} h'_{r-1} = 1_G$. On a donc $h_r = h'_r$, et par simplification on obtient

$$h_1 \cdots h_{r-1} = h'_1 \cdots h'_{r-1}.$$

Une récurrence facile montre alors l'égalité des décompositions. Cela achève la démonstration. □

2.6. Théorème. *Soit G un groupe, et soient $H_1, \ldots, H_r$ r sous-groupes de G (avec $r \geqslant 2$). Alors, $G = H_1 \odot \cdots \odot H_r$ si, et seulement si, les conditions suivantes sont satisfaites :*

(1) *le sous-groupe H_i est distingué dans G pour tout $i \in [\![1, r]\!]$;*

(2) *on a $G = \langle H_1, \ldots, H_r \rangle$;*

(3) *pour tout $i \in [\![1, r-1]\!]$, on a $H_1 \cdots H_i \cap H_{i+1} = \{1_G\}$.*

Démonstration. Si $G = H_1 \odot \cdots \odot H_r$, la proposition 2.5 montre que les conditions (2) et (3) sont vérifiées. De plus, on a vu au cours de la démonstration de cette proposition que la commutativité des éléments de H_i et H_j couplée à l'égalité $G = \langle H_1, \ldots, H_r \rangle$ impliquait que $H_1, \ldots, H_r$ sont des sous-groupes distingués de G.

Inversement, supposons les trois conditions du théorème vérifiées. D'après la proposition 2.5, il suffit de montrer que l'on a

$$h_i h_j = h_j h_i, \quad \text{pour tout } h_i \in H_i \text{ et tout } h_j \in H_j, \quad \text{dès que } i \neq j$$

pour pouvoir conclure. Soient $i \neq j$, $h_i \in H_i, h_j \in H_j$. On peut toujours supposer $i < j$. Posons $g = h_i h_j h_i^{-1} h_j^{-1}$. Le fait que H_i et H_j soient distingués dans G entraîne que $g \in H_i \cap H_j$. Or,

$$H_i \cap H_j \subset H_1 \cdots H_{j-1} \cap H_j = \{1_G\}.$$

Donc, $g = 1_G$, et ainsi $h_i h_j = h_j h_i$, d'où le résultat souhaité. □

2.7. Remarque. Au vu de la condition (1) sur les H_i, on pourrait remplacer (2) par la condition suivante :

(2′) on a $G = H_1 \cdots H_r$.

Néanmoins, il est plus facile a priori de vérifier (2) que (2′).

Dans le cas des groupes finis, on peut remplacer la condition (2) par une condition plus simple à vérifier.

2.8. Corollaire. *Soit G un groupe fini, et soient $H_1, \ldots, H_r$ des sous-groupes de G, en nombre égal à r (avec $r \geqslant 2$). Alors, $G = H_1 \odot \cdots \odot H_r$ si, et seulement si, les conditions suivantes sont satisfaites :*

(1) *le sous-groupe H_i est distingué dans G, pour tout $i \in [\![1, r]\!]$;*

(2) *on a $|G| = |H_1| \times \cdots \times |H_r|$;*

(3) *pour tout $i \in [\![1, r-1]\!]$, on a $H_1 \cdots H_i \cap H_{i+1} = \{1_G\}$.*

Démonstration. Si $G = H_1 \odot \cdots \odot H_r$, alors $G \simeq H_1 \times \cdots \times H_r$ par la remarque 2.4, d'où la condition (2). Les deux autres conditions proviennent des résultats précédents.

Pour montrer la réciproque, supposons tout d'abord que $r = 2$. On sait déjà que les sous-groupes H_1 et H_2 sont distingués dans G, d'intersection triviale. D'après le théorème 2.6, il reste à démontrer que $G = \langle H_1, H_2 \rangle$ pour pouvoir conclure. Comme H_1 est distingué dans G, cela revient à vérifier l'égalité $G = H_1 H_2$, par le lemme 1.1. Or, les conditions (2) et (3) impliquent que $|G| = |H_1 H_2|$, d'après le lemme 1.2. Comme $H_1 H_2 \subset G$, on en déduit $G = H_1 H_2$, d'où le résultat.

Supposons maintenant que $r \geqslant 3$, et posons $H = \langle H_1, \ldots, H_{r-1} \rangle$. Comme les sous-groupes $H_1, \ldots, H_{r-1}$ sont distingués dans G, ils sont distingués dans H. Puisque $H_1 \cdots H_i \cap H_{i+1} = \{1\}$ pour tout $i \in [\![1, r-2]\!]$, le théorème 2.6 montre que l'on a l'égalité

$$H = H_1 \odot \cdots \odot H_{r-1}.$$

En particulier, on a $H = H_1 \cdots H_{r-1}$, ainsi que

$$H \simeq H_1 \times \cdots \times H_{r-1}.$$

Par conséquent, $|H| = |H_1| \times \cdots \times |H_{r-1}|$. La condition (2) se récrit alors

$$|G| = |H| \times |H_r|.$$

De plus, la condition (3) appliquée à $i = r-1$ montre que $H \cap H_r = \{1_G\}$. Enfin, le fait que $H_1, \ldots, H_r$ soient distingués dans G impliquent aisément que H et H_r sont distingués dans G. Le cas $r = 2$ montre alors que l'on a

$$G = H \odot H_r = (H_1 \odot \cdots \odot H_{r-1}) \odot H_r = H_1 \odot \cdots \odot H_r,$$

la dernière égalité découlant facilement des définitions. Cela achève la démonstration. □

2.9. Corollaire. *Soit G un groupe fini, et soient H et K deux sous-groupes de G. Alors, on a $G = H \odot K$ si, et seulement si, les conditions suivantes sont satisfaites :*

(1) *les sous-groupes H et K sont distingués dans G (ou alternativement : on a $hk = hk$, pour tout $h \in H$, et tout $k \in K$) ;*

(2) *on a $|G| = |H| \times |K|$;*

(3) *on a $H \cap K = \{1_G\}$.*

Signalons, pour finir, que la notion de produit direct externe et interne se généralise aux familles quelconques de groupes.

2.10. Définition. Soit $(G_i)_{i\in I}$ une famille de groupes. Le *produit* des groupes G_i, $i \in I$ est le groupe

$$\prod_{i\in I} G_i = \{(x_i)_{i\in I} \mid x_i \in G_i\},$$

le produit de deux éléments étant fait composante par composante.

Le *coproduit* ou la *somme directe externe* des groupes G_i, $i \in I$ est le sous-groupe $\coprod_{i\in I} G_i$ des éléments $(x_i)_{i\in I} \in \prod_{i\in I} G_i$, tels que $x_i = 1_{G_i}$ pour presque tout $i \in I$.

Notation. Si $G_i = G$, pour tout $i \in I$, les groupes $\prod_{i\in I} G_i$ et $\coprod_{i\in I} G_i$ sont notés G^I et $G^{(I)}$ respectivement. Ce sont donc respectivement l'ensemble des familles d'éléments de G, et l'ensemble des familles d'éléments de G tous égaux à 1_G sauf un nombre fini.

2.11. Définition. Soit G un groupe, et soient $(H_i)_{i\in I}$ une famille de sous-groupes de G. On dit que G est le *coproduit direct interne* ou la *somme directe interne* de la famille de sous-groupes $(H_i)_{i\in I}$ si les deux conditions suivantes sont réalisées :

(1) on a $h_i h_j = h_j h_i$, pour tous $h_i \in H_i$, $h_j \in H_j$ et tous $i, j \in I$, $i \neq j$.

(2) tout élément $g \in G$ s'écrit de manière **unique** sous la forme

$$g = \prod_{i\in I} h_i, \quad h_i \in H_i,$$

où les h_i sont presque tous égaux à 1_G.

On le note $G = \bigodot_{i\in I} H_i$.

Bien entendu, si I est fini, on retrouve la notion usuelle de produit direct interne d'un nombre fini de sous-groupes.

2.12. Remarque. Si $G = \bigodot_{i\in I} H_i$, on a $G \simeq \coprod_{i\in I} H_i$.

3. Produits semi-directs

Soit G un groupe, et soient H et K deux sous-groupes de G. Reprenons l'étude de HK. On a vu dans le premier paragraphe que si H est distingué dans G, alors $\langle H, K\rangle = HK$.

En particulier, HK est un sous-groupe de G. On devrait pouvoir ainsi écrire le produit de deux éléments hk, $h'k' \in HK$ sous la forme $h''k''$.

C'est possible en utilisant la distinction de H. En effet, on a

$$(hk)(h'k') = \big(h(kh'k^{-1})\big)kk' \in HK.$$

Bien sûr, un élément de HK peut avoir plusieurs écritures différentes. L'écriture sera unique si, et seulement si, $H \cap K = \{1_G\}$.

En effet, si l'écriture est unique, alors pour tout $g \in H \cap K$, on a

$$g1_G = 1_G g \in HK,$$

et donc $g = 1_G$. Inversement, si $H \cap K = \{1_G\}$, alors pour tous h, $h' \in H$, et tous k, $k' \in K$, on a

$$hk = h'k' \implies (h')^{-1}h = k'k^{-1},$$

et donc $k'k^{-1} \in H \cap K$. Ainsi, $k'k^{-1} = 1_G$, soit $k = k'$, et l'on déduit ensuite que $h = h'$.

En particulier, si $H \cap K = \{1_G\}$, on a une bijection

$$\begin{aligned} H \times K &\longrightarrow HK \\ (h,k) &\longmapsto hk, \end{aligned}$$

mais cette bijection n'est **pas** un morphisme de groupes en général. En fait, le groupe HK est isomorphe à $H \times K$, mais muni de la loi de groupe

$$(h,k) * (h',k') = \big(h(kh'k^{-1}), kk'\big).$$

Remarquons maintenant que l'application

$$\begin{aligned} \theta : K &\longrightarrow \mathrm{Aut}(H) \\ k &\longmapsto \theta_k = \mathrm{Int}(k)_{|H} \end{aligned}$$

est un morphisme de groupes et que l'on a

$$(h,k) * (h',k') = \big(h\theta_k(h'), kk'\big).$$

Plus généralement, on a le résultat suivant.

3.1. Proposition. *Soient H et K deux groupes, et soit*

$$\begin{aligned} \rho : K &\longrightarrow \mathrm{Aut}(H) \\ k &\longmapsto \rho_k \end{aligned}$$

un morphisme de groupes. Alors, la loi interne

$$\begin{aligned} (H \times K) \times (H \times K) &\longrightarrow H \times K \\ \big((h,k),(h',k')\big) &\longmapsto (h,k) * (h',k') = \big(h\rho_k(h'), kk'\big) \end{aligned}$$

est une loi de groupe, de neutre $(1_H, 1_K)$.

Démonstration. Montrons que $(1_H, 1_K)$ est un élément neutre. On a

$$(h,k) * (1_H, 1_K) = \big(h\rho_k(1_H), k\big) = (h,k), \quad \text{pour tout } h \in H,\ k \in K,$$

car ρ_k est un élément de $\mathrm{Aut}(H)$, donc en particulier un morphisme de groupes. On a aussi

$$(1_H,1_K)*(h',k') = \big(\rho_{1_K}(h'),k'\big) = (h',k'), \quad \text{pour tout } h' \in H \text{ et tout } k' \in K,$$

car ρ étant un morphisme de groupes, il envoie 1_K sur Id_H. Donc, $(1_H, 1_K)$ est un élément neutre pour la loi interne.

Soit $(h,k) \in H \times K$, et montrons que $(\rho_k^{-1}(h^{-1}), k^{-1})$ est un inverse de (h,k).

On a

$$(h,k) * (\rho_k^{-1}(h^{-1}), k^{-1}) = \big(h\rho_k\big(\rho_k^{-1}(h^{-1})\big), 1_K\big) = (hh^{-1}, 1_K) = (1_H, 1_K).$$

On a aussi

$$\big(\rho_k^{-1}(h^{-1}), k^{-1}\big) * (h,k) = \big(\rho_k^{-1}(h^{-1})\rho_{k^{-1}}(h), 1_K\big).$$

Comme ρ est un morphisme de groupes, on a $\rho_{k^{-1}} = \rho_k^{-1}$, et l'on obtient donc

$$\begin{aligned} \big(\rho_k^{-1}(h^{-1}), k^{-1}\big) * (h,k) &= \big(\rho_{k^{-1}}(h^{-1})\rho_{k^{-1}}(h), 1_K\big) \\ &= \big(\rho_{k^{-1}}(h^{-1}h), 1_K\big) \\ &= \big(\rho_{k^{-1}}(1_H), 1_K\big) = (1_H, 1_K), \end{aligned}$$

les deux dernières égalités provenant du fait que $\rho_{k^{-1}}$ est un morphisme de groupes. Il reste à vérifier l'associativité. On a d'une part

$$\begin{aligned} \big((h,k) * (h',k')\big) * (h'',k'') &= \big(h\rho_k(h'), kk'\big) * (h'',k'') \\ &= \big(h\rho_k(h')\rho_{kk'}(h''), (kk')k''\big), \end{aligned}$$

et d'autre part

$$\begin{aligned} (h,k) * \big((h',k') * (h'',k'')\big) &= (h,k) * \big(h'\rho_{k'}(h''), k'k''\big) \\ &= \big(h\rho_k\big(h'\rho_{k'}(h'')\big), k(k'k'')\big). \end{aligned}$$

Puisque K est un groupe, on a bien sûr $(kk')k'' = k(k'k'')$. Maintenant, ρ étant un morphisme de groupes, on a $\rho_{kk'} = \rho_k \circ \rho_{k'}$, et donc

$$h\rho_k(h')\rho_{kk'}(h'') = h\rho_k(h')\rho_k\big(\rho_{k'}(h'')\big).$$

Comme ρ_k est un morphisme de groupes, on a alors

$$h\rho_k(h')\rho_{kk'}(h'') = h\rho_k\big(h'\rho_{k'}(h'')\big).$$

Cela achève la démonstration. □

3.2. Définition. Soient H et K deux groupes, et soit

$$\rho : K \longrightarrow \mathrm{Aut}(H)$$

un morphisme de groupes. L'ensemble $H \times K$ muni de la loi de groupe précédente est appelé *produit semi-direct externe* de H et K par ρ, et est noté $H \times_\rho K$.

3.3. Remarque. On vérifie aisément que le sous-groupe $H' = \{(h, 1_K) \mid h \in H\}$ est distingué dans $H \times_\rho K$, et isomorphe à H.

3.4. Lemme. *Soient H et K deux groupes, et soit*

$$\begin{aligned} \rho : K &\longrightarrow \mathrm{Aut}(H) \\ k &\longmapsto \quad \rho_k \end{aligned}$$

un morphisme de groupes. Les propriétés suivantes sont équivalentes :

(1) *on a $H \times_\rho K = H \times K$;*

(2) *le morphisme $\rho : K \longrightarrow \mathrm{Aut}(H)$ est trivial ;*

(3) *le sous-groupe $K' = \{(1_H, k) \mid k \in K\}$ est distingué dans $H \times_\rho K$.*

Démonstration

(3) $\Longrightarrow$ (2). Soient $h \in H$, $k \in K$. On a

$$(h, 1_K) * (1_H, k) * (h, 1_K)^{-1} = (h, k) * (h, 1_K)^{-1} = (h, k) * (h^{-1}, 1_K),$$

soit encore

$$(h, 1_K) * (1_H, k) * (h, 1_K)^{-1} = \big(h\rho_k(h^{-1}), k\big).$$

Puisque K' est distingué dans $H \times_\rho K$, on en déduit que

$$h\rho_k(h^{-1}) = h\rho_k(h)^{-1} = 1_H, \quad \text{pour tout } h \in H \text{ et tout } k \in K.$$

Autrement dit, on a $\rho_k(h) = h$, pour tout $h \in H$ et tout $k \in K$, i.e. ρ est trivial.

(2) $\Longrightarrow$ (1). Évident.

(1) $\Longrightarrow$ (3). Découle des propriétés du produit direct. □

3.5. Remarque. Ce résultat devient faux si l'on remplace l'égalité par un isomorphisme dans le point (1) (cf. exercice 4).

On s'intéresse maintenant au produit semi-direct interne.

3.6. Définition. Soit G un groupe, et soient H et K deux sous-groupes de G. On dit que G est le *produit semi-direct interne* de H et K si les conditions suivantes sont satisfaites :

(1) le sous-groupe H est distingué dans G ;

(2) tout élément $g \in G$ s'écrit de manière unique $g = hk$, $h \in H$, $k \in K$;

On le note $G = H \rtimes K$.

3.7. Remarque. Soit G un groupe, et soient H et K deux sous-groupes, avec H distingué dans G. Considérons le morphisme de groupes

$$\begin{aligned} \theta : K &\longrightarrow \mathrm{Aut}(H) \\ k &\longmapsto \mathrm{Int}(k). \end{aligned}$$

Alors, $G = H \rtimes K$ si, et seulement si, l'application

$$\begin{aligned} H \times_\theta K &\longrightarrow G \\ (h, k) &\longmapsto hk \end{aligned}$$

est un isomorphisme de groupes.

En particulier, un produit semi-direct interne est isomorphe à un certain produit semi-direct externe.

3.8. Exemples. Voici quelques exemples de produits semi-directs internes :

(1) $\mathfrak{S}(E) = \mathfrak{A}(E) \rtimes \langle \tau \rangle$, où τ est une transposition fixée ;

(2) $\mathrm{GL}_n(K) = \mathrm{SL}_n(K) \rtimes \mathrm{D_n(K)}$, où $\mathrm{D_n(K)}$ est le sous-groupe de $\mathrm{GL}_n(K)$ formé des matrices de dilatation et de la matrice identité ;

(3) soit $\sigma \in D_n$ une rotation de mesure d'angle $\frac{2\pi}{n}$ et soit $\tau \in D_n$ une symétrie orthogonale dont l'axe passe par le centre du n-gone et un des sommets. Posons $H = \langle \sigma \rangle$ et $K = \langle \tau \rangle$. Alors, $D_n = H \rtimes K$.

On a dans ce cas $D_n \simeq H \times_\rho K$, où $\rho_1 = \mathrm{Id}_H$ et ρ_τ est le passage à l'inverse.

Les vérifications sont laissées au lecteur.

3.9. Proposition. *Soit G un groupe, et soient H et K deux sous-groupes. Alors, $G = H \rtimes K$ si, et seulement si, les conditions suivantes sont vérifiées :*

(1) *le sous-groupe H est distingué dans G ;*

(2) *on a $G = \langle H, K \rangle$ (ou, de manière équivalente, $G = HK$) ;*

(3) *on a $H \cap K = \{1_G\}$.*

Démonstration. Supposons que $G = H \rtimes K$. Alors, par définition H est distingué dans G. De plus, comme tout élément de G est produit d'un élément de H et d'un élément de K, on a $G = HK$ et donc $G = \langle H, K \rangle$. On a déjà constaté que l'unicité d'écriture dans HK implique (3).

Inversement, les conditions (1) et (2) impliquent les égalités

$$G = HK = \langle H, K \rangle,$$

par le lemme 1.1. La condition $H \cap K = \{1_G\}$ implique l'unicité de l'écriture, comme on l'a déjà remarqué. □

3.10. Remarque. Si $G = H \rtimes K$, alors $G = H \odot K$ si, et seulement si, K est distingué dans G.

En effet, le sens direct est clair. Réciproquement, si K est un sous-groupe distingué de G, on utilise la proposition précédente et le théorème 2.6 pour conclure.

On continue en donnant des conditions nécessaires et suffisantes pour avoir un produit semi-direct interne dans le cas des groupes finis.

3.11. Corollaire. *Soit G un groupe fini, et soient H et K deux sous-groupes. Alors, $G = H \rtimes K$ si, et seulement si, les conditions suivantes sont vérifiées :*

(1) *le sous-groupe H est distingué dans G ;*
(2) *on a $|G| = |H| \times |K|$;*
(3) *on a $H \cap K = \{1_G\}$.*

Démonstration. Similaire à celle du corollaire 2.9. □

Nous achevons ce chapitre en donnant d'autres critères, basés sur la notion de suite exacte.

3.12. Définition. Une *suite exacte* de groupes, notée

$$\{1\} \longrightarrow H \xrightarrow{\iota} G \xrightarrow{\pi} K \longrightarrow \{1\},$$

est la donnée d'un morphisme de groupes injectif $\iota : H \longrightarrow G$, et d'un morphisme de groupes surjectif $\pi : G \longrightarrow K$ tels que $\mathrm{Im}(\iota) = \mathrm{Ker}(\pi)$.

On dit que la suite exacte

$$\{1\} \longrightarrow H \xrightarrow{\iota} G \underset{s}{\overset{\pi}{\rightleftarrows}} K \longrightarrow \{1\}$$

est *scindée* s'il existe un morphisme de groupes $s : K \longrightarrow G$ tel que

$$\pi \circ s = \mathrm{Id}_K.$$

Un tel morphisme s est appelé une *section* de π, ou encore une section de la suite exacte.

3.13. Remarque. Si la suite exacte

$$\{1\} \longrightarrow H \xrightarrow{\iota} G \underset{s}{\overset{\pi}{\rightleftarrows}} K \longrightarrow \{1\}$$

est scindée, alors G contient un sous-groupe isomorphe à K.

En effet, si $s : K \longrightarrow G$ vérifie $\pi \circ s = \mathrm{Id}_K$, alors s est injective, et donc induit un isomorphisme de $K \simeq \mathrm{Im}(s)$. Ainsi, $\mathrm{Im}(s)$ est le sous-groupe recherché.

3.14. Exemples

(1) Soit G un groupe, soit H un sous-groupe distingué de G. Si $\iota : H \longrightarrow G$ est l'inclusion et $\pi : G \longrightarrow G/H$ est la projection canonique, alors la suite

$$\{1\} \longrightarrow H \xrightarrow{\iota} G \xrightarrow{\pi} K \longrightarrow \{1\}$$

est une suite exacte. Celle-ci n'est pas scindée en général.
Par exemple, si $G = \mathbb{Z}$ et $H = 2\mathbb{Z}$,

$$\{1\} \longrightarrow 2\mathbb{Z} \hookrightarrow \mathbb{Z} \longrightarrow \mathbb{Z}/2\mathbb{Z} \longrightarrow \{1\} ;$$

cette suite exacte n'est pas scindée. Dans le cas contraire, $\mathbb{Z}$ contiendrait un sous-groupe isomorphe à $\mathbb{Z}/2\mathbb{Z}$ par la remarque précédente, donc un élément d'ordre 2, ce qui est absurde.

(2) Soit $G = H \times_\rho K$, où $\rho : K \longrightarrow \mathrm{Aut}(H)$ est un morphisme de groupes. On considère les applications

$$\begin{array}{rcl} \iota_1 : H & \longrightarrow & G \\ h & \longmapsto & (h,1) \end{array} \quad \text{et} \quad \begin{array}{rcl} \pi_2 : G & \longrightarrow & K \\ (h,k) & \longmapsto & k. \end{array}$$

Alors, on vérifie en utilisant les définitions que ι_1 et π_2 sont des morphismes de groupes, et que la suite

$$\{1\} \longrightarrow H \xrightarrow{\iota_1} G \underset{\iota_2}{\overset{\pi_2}{\rightleftarrows}} K \longrightarrow \{1\}$$

est une suite exacte scindée par le morphisme

$$\begin{array}{rcl} \iota_2 : K & \longrightarrow & G \\ k & \longmapsto & (1,k). \end{array}$$

La proposition précédente caractérise les produits semi-directs en termes de suite exacte.

3.15. Théorème. *Soient G, H, K des groupes. Alors, $G \simeq H \times_\rho K$ (pour un certain morphisme $\rho : K \longrightarrow \mathrm{Aut}(H)$) si, et seulement si, il existe une suite exacte scindée*

$$\{1\} \longrightarrow H \xrightarrow{\iota} G \underset{s}{\overset{\pi}{\rightleftarrows}} K \longrightarrow \{1\}.$$

Dans ce cas, si $s : K \longrightarrow G$ est une section de π, on a

$$G = \mathrm{Im}(\iota) \rtimes \mathrm{Im}(s) \simeq H \times_\rho K,$$

où, pour tout $h \in H$ et tout $k \in K$, $\rho_k(h)$ est l'unique élément de H vérifiant

$$s(k)\iota(h)s(k)^{-1} = \iota\big(\rho_k(h)\big).$$

Démonstration. Supposons que $G \simeq H \times_\rho K$, et soit $\varphi : H \times_\rho K \xrightarrow{\sim} G$ un isomorphisme. Considérons les morphismes

$$\iota = \varphi \circ \iota_1, \quad s = \varphi \circ \iota_2 \quad \text{et} \quad \pi = \pi_2 \circ \varphi^{-1},$$

où ι_1, ι_2 et π_2 sont définis dans l'exemple 3.14 (2).

$$\begin{array}{ccccccccc}
\{1\} & \longrightarrow & H & \xrightarrow{\iota_1} & H \times_\rho K & \underset{\iota_2}{\overset{\pi_2}{\rightleftarrows}} & K & \longrightarrow & \{1\} \\
 & & \| & & \wr\downarrow \varphi & & \| & & \\
\{1\} & \longrightarrow & H & \xrightarrow{\iota = \varphi \circ \iota_1} & G & \underset{s = \varphi \circ \iota_2}{\overset{\pi = \pi_2 \circ \varphi^{-1}}{\rightleftarrows}} & K & \longrightarrow & \{1\}
\end{array}$$

Alors, ι est injective, π est surjective, et l'on a

$$\mathrm{Ker}(\pi) = \varphi\big(\mathrm{Ker}(\pi_2)\big) = \varphi\big(\mathrm{Im}(\iota_1)\big) = \mathrm{Im}(\iota).$$

On a donc une suite exacte

$$\{1\} \longrightarrow H \xrightarrow{\iota} G \underset{s}{\overset{\pi}{\rightleftarrows}} K \longrightarrow \{1\}.$$

De plus, elle est scindée puisque

$$\pi \circ s = (\pi_2 \circ \varphi^{-1}) \circ (\varphi \circ \iota_2) = \pi_2 \circ \iota_2 = \mathrm{Id}_K,$$

toujours d'après le même exemple.

Inversement, supposons que l'on ait une suite exacte

$$\{1\} \longrightarrow H \xrightarrow{\iota} G \underset{s}{\overset{\pi}{\rightleftarrows}} K \longrightarrow \{1\}$$

scindée par $s : K \longrightarrow G$.

Vérifions tout d'abord que $G = \mathrm{Im}(\iota) \rtimes \mathrm{Im}(s)$. Puisque $\mathrm{Im}(\iota) = \mathrm{Ker}(\pi)$, le sous-groupe $\mathrm{Im}(\iota)$ est distingué dans G. Soit $x \in \mathrm{Im}(\iota) \cap \mathrm{Im}(s)$. Il existe donc $h \in H$ et $k \in K$ tel que $x = \iota(h) = s(k)$. On a donc

$$\pi(x) = \pi\big(\iota(h)\big) = 1_K = \pi\big(s(k)\big) = k,$$

d'où $k = 1_K$, et ainsi $x = s(1_K) = 1_G$. On a donc $\mathrm{Im}(\iota) \cap \mathrm{Im}(s) = \{1_G\}$. Enfin, soit $x \in G$. Alors, $k = \pi(x) \in K$. Vérifions que $xs(k)^{-1} \in \mathrm{Im}(\iota)$. Pour cela, il suffit de montrer que $xs(k)^{-1} \in \mathrm{Ker}(\pi)$, puisque la suite est exacte. Or, on a

$$\pi(xs(k)^{-1}) = \pi\big(xs(k^{-1})\big) = \pi(x)\pi\big(s(k^{-1})\big) = \pi(x)k^{-1} = 1_K,$$

d'où le résultat. Ainsi, il existe $h \in H$ tel que $xs(k)^{-1} = \iota(h)$, et par conséquent, $x = \iota(h)s(k) \in \mathrm{Im}(\iota)\mathrm{Im}(s)$. Par la proposition 3.9, on a alors

$$G = \mathrm{Im}(\iota) \rtimes \mathrm{Im}(s).$$

Enfin, puisque $\mathrm{Im}(\iota)$ est distingué dans G, remarquons que pour tout $h \in K$,

et tout $k \in K$, il existe un élément $h' \in H$ tel que

$$s(k)\iota(h)s(k)^{-1} = \iota(h').$$

Un tel élément h' est unique par injectivité de ι. On le note $\rho_k(h)$. Remarquons maintenant que pour tous k, $k' \in K$, on a

$$\rho_{kk'} = \rho_k \circ \rho_{k'}.$$

En effet, par définition, $\rho_k : H \longrightarrow H$ est l'unique application vérifiant

$$\mathrm{Int}_{s(k)} \circ \iota = \iota \circ \rho_k,$$

et $\rho_{k'} : H \longrightarrow H$ est l'unique application vérifiant

$$\mathrm{Int}_{s(k')} \circ \iota = \iota \circ \rho_{k'}.$$

On a alors

$$\begin{aligned} \mathrm{Int}_{s(kk')} \circ \iota &= \mathrm{Int}_{s(k)s(k')} \circ \iota \\ &= \mathrm{Int}_{s(k)} \circ (\mathrm{Int}_{s(k')} \circ \iota) \\ &= (\mathrm{Int}_{s(k)} \circ \iota) \circ \rho_{k'} \\ &= \iota \circ (\rho_k \circ \rho_{k'}). \end{aligned}$$

Comme $\rho_{kk'} : H \longrightarrow H$ est l'unique application vérifiant

$$\mathrm{Int}_{s(kk')} \circ \iota = \iota \circ \rho_{kk'},$$

on obtient bien l'égalité désirée. En particulier, pour tout $k \in K$, on a

$$\rho_k \circ \rho_{k^{-1}} = \rho_{k^{-1}} \circ \rho_k = \rho_{1_K} = \mathrm{Id}_H,$$

la dernière égalité découlant de la définition de ρ_{1_K}. Ainsi, $\rho_k \in \mathrm{Aut}(H)$ pour tout $k \in K$, et l'égalité démontrée précédemment entraîne alors que l'application $\rho : K \longrightarrow \mathrm{Aut}(H)$ est un morphisme de groupes. On peut donc former le groupe $H \times_\rho K$. Il reste donc à voir que $G \simeq H \times_\rho K$. Pour cela, montrons que l'application

$$\begin{aligned} \varphi : H \times_\rho K &\longrightarrow G \\ (h,k) &\longmapsto \iota(h)s(k) \end{aligned}$$

est un isomorphisme de groupes. Pour tous h, $h' \in H$ et pour tous k, $k' \in K$, on a

$$\begin{aligned} \varphi\big((h,k)(h',k')\big) &= \varphi\big(h\rho_k(h'), kk'\big) \\ &= \iota\big(h\rho_k(h')\big)s(kk') \\ &= \iota(h)\iota\big(\rho_k(h')\big)s(k)s(k') \\ &= \iota(h)\big(s(k)\iota(h')s(k)^{-1}\big)s(k)s(k') \\ &= \big(\iota(h)s(k)\big)\big(\iota(h')s(k')\big) \\ &= \varphi\big((h,k)\big)\varphi\big((h',k')\big). \end{aligned}$$

Ainsi, φ est un morphisme de groupes. Comme $G = \mathrm{Im}(\iota) \rtimes \mathrm{Im}(s)$, φ est surjectif. L'injectivité découle du fait que $\mathrm{Im}(\iota) \cap \mathrm{Im}(s) = \{1_G\}$. Par conséquent, φ est bien un isomorphisme. Cela achève la démonstration. □

3.16. Corollaire. *Soient G, H, K des groupes abéliens. Alors, $G \simeq H \times K$ si, et seulement si, il existe une suite exacte scindée*

$$\{1\} \longrightarrow H \xrightarrow{\iota} G \underset{s}{\overset{\pi}{\rightleftarrows}} K \longrightarrow \{1\}.$$

Dans ce cas, si $s : K \longrightarrow G$ est une section, on a

$$G = \mathrm{Im}(\iota) \odot \mathrm{Im}(s).$$

Démonstration. Le sens direct est obtenu en prenant pour ρ le morphisme trivial dans le théorème précédent. Pour la réciproque, on remarque que les sous-groupes $\mathrm{Im}(\iota)$ et $\mathrm{Im}(s)$ sont distingués dans G. Comme tout élément de G s'écrit de manière unique comme le produit d'un élément de $\mathrm{Im}(\iota)$ et d'un élément de $\mathrm{Im}(s)$ par ce même théorème, on a bien $G = \mathrm{Im}(\iota) \odot \mathrm{Im}(s)$. Mais alors,

$$G \simeq \mathrm{Im}(\iota) \odot \mathrm{Im}(s) \simeq H \times K,$$

ce qui achève la démonstration. □

3.17. Remarque. Si $G \simeq H \times_\rho K$, le théorème 3.15 montre qu'il existe deux sous-groupes G_1, G_2 de G, isomorphes à H et K respectivement, et tels que $G = G_1 \rtimes G_2$.

Ainsi, un produit semi-direct externe est un produit semi-direct interne (pour d'autres sous-groupes) et vice-versa. La même chose est vraie pour les produits directs. L'intérêt de la version externe pour les produits directs ou semi-directs est le même que pour les sommes directes d'espaces vectoriels : celui de pouvoir construire de nouveaux groupes à partir de groupes abstraits ne vivant pas dans un même groupe. La version interne est là pour décortiquer la structure d'un groupe donné.

On peut aussi énoncer la version interne de ces théorèmes, qui découle directement des résultats précédents, mais qui peut aussi se démontrer directement.

3.18. Corollaire. *Soit G un groupe, et soient H et K deux sous-groupes de G. Alors, on a $G = H \rtimes K$ (resp. $G = H \odot K$ si G est abélien) si, et seulement si, il existe un morphisme $\pi : G \longrightarrow K$ surjectif, de noyau H, tel que $\pi_{|K} = \mathrm{Id}_K$.*

4. Exercices

E1. Soit G un groupe, et soient $H_1, \ldots, H_r$ des sous-groupes de G. Montrer que $G = H_1 \odot \cdots \odot H_r$ si, et seulement si, l'application

$$\begin{aligned} H_1 \times \cdots \times H_r &\longrightarrow G \\ (h_1, \ldots, h_r) &\longmapsto h_1 \cdots h_r. \end{aligned}$$

est un isomorphisme de groupes.

E2. Soit E un ensemble fini, et soient σ, $\sigma' \in \mathfrak{S}(E)$ deux permutations à supports disjoints. Montrer l'égalité

$$\langle \sigma, \sigma' \rangle = \langle \sigma \rangle \odot \langle \sigma' \rangle.$$

E3. Soient $p_1, \ldots, p_k$ des nombres premiers distincts, et soit G le sous-groupe de $\mathbb{Q}^\times$ engendré par $p_1, \ldots, p_k$.

a. Montrer que $G \simeq \mathbb{Z}^k$.

b. Montrer que $G = \langle p_1 \rangle \odot \langle p_1 p_2 \rangle \odot \langle p_2 p_3 \rangle \odot \cdots \odot \langle p_{k-1} p_k \rangle$.

E4. Soit $\rho : \mathbb{Z}/2\mathbb{Z} \longrightarrow \mathrm{Aut}(\mathfrak{S}_3)$ l'unique morphisme tel que $\rho_{\overline{1}} = \mathrm{Int}\big((12)\big)$. Montrer que $\rho \neq 1$, mais que $\mathfrak{S}_3 \times_\rho \mathbb{Z}/2\mathbb{Z} \simeq \mathfrak{S}_3 \times \mathbb{Z}/2\mathbb{Z}$.

Ainsi, on ne peut pas remplacer « $=$ » par « $\simeq$ » dans le lemme 3.4.

E5. Montrer que le groupe des quaternions n'est pas isomorphe à un produit semi-direct de deux groupes non triviaux.

E6. Détailler les exemples du point 3.8.

E7. Soit $G = H \rtimes K$, et soit N un sous-groupe de G contenant H. Montrer que $N = H \rtimes (N \cap K)$.

E8.

a. Soient H, H', K, K' des groupes, et soit $\theta : K \longrightarrow \mathrm{Aut}(H)$ un morphisme. Si $\varphi : H \xrightarrow{\sim} H'$ et $\psi : K' \xrightarrow{\sim} K$ sont des isomorphismes, montrer que l'on a

$$H \times_\theta K \simeq H' \times_{\theta'} K',$$

où $\theta' : K' \longrightarrow \mathrm{Aut}(H')$ est le morphisme défini par

$$\theta'_{k'} = \varphi \circ \theta_{\psi(k')} \circ \varphi^{-1}, \quad \text{pour tout } k' \in K'.$$

b. Soient ρ_1, $\rho_2 : K \longrightarrow \mathrm{Aut}(H)$. Montrer que $H \times_{\rho_1} K \simeq H \times_{\rho_2} K$ dans chacun des cas suivants.

(1) Il existe $\alpha \in \mathrm{Aut}(K)$ tel que $\rho_2 = \rho_1 \circ \alpha$.

(2) Il existe $u \in \mathrm{Aut}(H)$ tel que $\rho_{2,k} = u \circ \rho_{1,k} \circ u^{-1}$, pour tout $k \in K$.

c. On suppose que K est cyclique d'ordre n, et que $\rho_1(K) = \rho_2(K)$. On veut montrer que $H \times_{\rho_1} K \simeq H \times_{\rho_2} K$. Soit k_0 un générateur de K, et soit $m = o(\rho_{1,k_0})$.

(i) Montrer qu'il existe $d \geqslant 1$ premier à m tel que $\rho_{2,k_0} = \rho_{1,k_0}^d$.

(ii) Justifier que $m \mid n$. Soit m' le produit des diviseurs premiers de n ne divisant pas m. Montrer qu'il existe $\ell \in \mathbb{Z}$ tel que

$$\begin{cases} \ell & \equiv & d \ [m] \\ \ell & \equiv & 1 \ [m'], \end{cases}$$

et montrer que ℓ est premier à n.

(iii) En déduire qu'il existe $\alpha \in \mathrm{Aut}(K)$ tel que $\rho_2 = \rho_1 \circ \alpha$ et conclure.

E9. Soit p un nombre premier impair, et soit G un groupe d'ordre $2p$.

a. Justifier que G admet un élément σ d'ordre p, ainsi qu'un élément τ d'ordre 2.

b. Montrer que $\langle\sigma\rangle$ est distingué dans G.

Indication. Que vaut $[G : \langle\sigma\rangle]$?

c. Montrer que $G = \langle\sigma\rangle \rtimes \langle\tau\rangle$.

d. Soit $H = \langle\sigma\rangle$. On considère l'automorphisme

$$\begin{array}{rccc} \rho : & H & \longrightarrow & H \\ & h & \longmapsto & \tau h \tau^{-1}. \end{array}$$

(i) En utilisant l'exercice VI-37, montrer qu'il existe un entier $m \in \mathbb{Z}$ premier à p tel que

$$\rho(h) = h^m, \quad \text{pour tout } h \in H.$$

Montrer ensuite que $m \equiv \pm 1 \ [p]$.

(ii) Si $m \equiv 1 \ [p]$, montrer que G est cyclique d'ordre $2p$.

(iii) Si $m \equiv -1 \ [p]$, montrer que G est isomorphe à D_p.

e. Résumer le résultat obtenu.

E10. Trouver à isomorphisme près tous les produits semi-directs $H \times_\rho K$, où $|H| = 4$ et $|K| = 3$. Reconnaître $\mathfrak{A}_4$ parmi eux.

Indication. On pourra utiliser l'exercice 8.

E11. Soit K un corps de caractéristique 0 et soit $n \geqslant 2$. On note $\mathrm{Diag}_n(K)$ le sous-groupe de $\mathrm{GL}_n(K)$ dont les éléments sont les matrices diagonales inversibles. On pose

$$G = \{M \in \mathrm{GL}_n(K) \mid M\mathrm{Diag}_n(K)M^{-1} = \mathrm{Diag}_n(K)\}.$$

a. Vérifier que G est un sous-groupe de $\mathrm{GL}_n(K)$.

b. Montrer que $G \simeq (K^\times)^n \times_\rho \mathfrak{S}_n$, où ρ est un morphisme non trivial que l'on précisera.

E12. On note $\mathrm{GLA}(\mathbb{F}_5)$ le sous-ensemble de $\mathfrak{S}(\mathbb{F}_5)$ dont les éléments sont les bijections de la forme

$$\begin{aligned} \mathbb{F}_5 &\longrightarrow \mathbb{F}_5 \\ \overline{x} &\longmapsto \overline{a}\,\overline{x} + \overline{b}, \end{aligned}$$

avec $\overline{a} \in \mathbb{F}_5^\times$ et $\overline{b} \in \mathbb{F}_5$.

a. Vérifier que $\mathrm{GLA}(\mathbb{F}_5)$ est un sous-groupe de $\mathfrak{S}(\mathbb{F}_5)$ d'ordre 20, que $\overline{2}$ engendre $\mathbb{F}_5^\times$, et que les bijections

$$\begin{aligned} s : \mathbb{F}_5 &\longrightarrow \mathbb{F}_5 \\ \overline{x} &\longmapsto \overline{x} + \overline{1} \end{aligned} \quad \text{et} \quad \begin{aligned} t : \mathbb{F}_5 &\longrightarrow \mathbb{F}_5 \\ \overline{x} &\longmapsto \overline{2}\,\overline{x} \end{aligned}$$

engendrent $\mathrm{GLA}(\mathbb{F}_5)$.

b. Vérifier qu'il y a exactement trois morphismes de groupes

$$(\mathbb{Z}/5\mathbb{Z})^\times \longrightarrow \mathrm{Aut}(\mathbb{Z}/5\mathbb{Z})$$

non triviaux. Trouver tous les produits semi-directs $\mathbb{Z}/5\mathbb{Z} \times_\rho (\mathbb{Z}/5\mathbb{Z})^\times$ deux à deux non isomorphes.

Indication. On pourra utiliser les exercices VI-37 et 8, ainsi que regarder les éléments d'ordre 2.

c. Montrer que $\mathrm{GLA}(\mathbb{F}_5) = \langle s \rangle \rtimes \langle t \rangle \simeq \mathbb{Z}/5\mathbb{Z} \times_\rho (\mathbb{Z}/5\mathbb{Z})^\times$, où le morphisme de groupes $\rho : (\mathbb{Z}/5\mathbb{Z})^\times \longrightarrow \mathrm{Aut}(\mathbb{Z}/5\mathbb{Z})$ est à déterminer.

d. Exceptionnellement, on représente les éléments de $\mathbb{F}_5$ par des entiers dans $[\![1, 5]\!]$. Montrer alors que les images de s et t, via l'isomorphisme de groupes $\mathfrak{S}(\mathbb{F}_5) \simeq \mathfrak{S}_5$, sont

$$\sigma = (1\ 2\ 3\ 4\ 5) \quad \text{et} \quad \tau = (1\ 2\ 4\ 3).$$

E13. Cet exercice nécessite d'avoir étudié la partie du chapitre consacrée aux suites exactes.

a. Soit K un corps, et soit E un K-espace vectoriel de dimension $n \geqslant 1$.

(i) Montrer que l'on a une suite exacte scindée

$$\{1\} \longrightarrow \mathrm{SL}(E) \hookrightarrow \mathrm{GL}(E) \underset{s}{\overset{\det}{\rightleftarrows}} K^\times \longrightarrow \{1\}.$$

(ii) Montrer que $\mathrm{GL}(E) \simeq \mathrm{SL}(E) \times K^\times$ si, et seulement si, le morphisme

$$\begin{aligned} K^\times &\longrightarrow K^\times \\ x &\longmapsto x^n \end{aligned}$$

est un automorphisme.

b. Soit E un espace euclidien de dimension $n \geqslant 1$.

(i) Montrer que l'on a une suite exacte scindée

$$\{1\} \longrightarrow \mathrm{O}^+(E) \hookrightarrow \mathrm{O}(E) \underset{s}{\overset{\det}{\rightleftarrows}} \{\pm 1\} \longrightarrow \{1\}.$$

(ii) Déterminer les valeurs de n pour lesquelles on a un isomorphisme

$$\mathrm{O}(E) \simeq \mathrm{O}^+(E) \times \{\pm 1\}.$$

E14. Soit $G = H \times H$, où H est un groupe non abélien. On pose

$$H_1 = H \times \{1_H\} \quad \text{et} \quad H_2 = \{(h, h) \mid h \in H\}.$$

Montrer que

$$G = H_1 \rtimes H_2 \simeq H_1 \times_\rho H_2 \simeq H \times_{\rho'} H,$$

où ρ et ρ' sont à préciser.

Chapitre XI

Applications des groupes opérants à la structure des groupes finis

1. Théorème de Sylow

Que devient un p-Sylow lorsqu'il est devenu fou ? Un Sylow à grain.

Soit G un groupe fini. Comme nous l'avons vu auparavant, l'ordre de tout sous-groupe H divise l'ordre du groupe, mais étant donné un diviseur d de l'ordre de G, il n'existe pas nécessairement de sous-groupe de G d'ordre d. Grâce aux actions de groupes, nous allons voir que cela devient vrai lorsque d est une puissance d'un nombre premier.

1.1. Définition. Soit G un groupe fini et soit p un nombre premier.

Un *p-sous-groupe* de G est un sous-groupe de G qui est un p-groupe.

Écrivons $|G| = p^m q$, où $p \nmid q$ et $m \geqslant 0$. Un *p-sous-groupe de Sylow* de G, ou p-Sylow, est un p-sous-groupe de G d'ordre p^m. Autrement dit, un sous-groupe S est un p-Sylow du groupe G si, et seulement si, S est un p-groupe et $p \nmid [G : S]$.

Évidemment, rien ne garantit qu'un tel p-Sylow existe. Nous allons maintenant donner un exemple de groupes où un p-Sylow existe bel et bien.

1.2. Exemple. Soit $G = \mathrm{GL}_n(\mathbb{F}_p)$. C'est un groupe fini d'ordre

$$|G| = (p^n - 1)\times(p^n - p)\times\cdots\times(p^n - p^{n-1}).$$

En effet, pour qu'une matrice soit inversible, il faut et il suffit que ces colonnes soient linéairement indépendantes. On a donc $p^n - 1$ choix pour la première colonne (on ne peut pas prendre une colonne nulle). La deuxième colonne ne doit pas être multiple de la première. On a donc $p^n - p$ choix. La troisième ne doit pas être combinaison linéaire des deux premières, on a donc $p^n - p^2$ choix, etc.

On a donc

$$|G| = (p^n - 1)\times\cdots\times(p^n - p^{n-1}) = p^{1+2+\cdots+(n-1)}q = p^{n(n-1)/2}q, \quad p \nmid q.$$

Soit S le sous-groupe des matrices triangulaires supérieures avec des $\overline{1}$ sur la diagonale. Autrement dit,

$$S = \{(a_{ij}) \in G \mid a_{ij} = \overline{0}, \text{ si } i > j,\ a_{ii} = \overline{1}\}.$$

Alors, S est un sous-groupe de G d'ordre $p^{n(n-1)/2}$, donc un p-Sylow.

Nous allons maintenant montrer l'existence d'un p-Sylow pour tout groupe G. On rappelle que si H un sous-groupe, et si $g \in G$, alors

$$gHg^{-1} = \{ghg^{-1} \mid g \in G\}.$$

C'est un sous-groupe de G, puisque c'est l'image de H par $\mathrm{Int}(g)$. De plus, si H est fini, alors il en est de même de gHg^{-1}, et on a

$$|gHg^{-1}| = |H|,$$

puisque Int_g est un automorphisme.

Si H et H' deux sous-groupes, on rappelle que H et H' sont *conjugués* s'il existe $g \in G$ tel que $H' = gHg^{-1}$. On dit aussi que H' est un conjugué de H (cf. définition VI-1.14).

1.3. Remarque. Soit G un groupe fini et soit p un nombre premier. Alors, tout conjugué d'un p-Sylow de G est encore un p-Sylow de G (cela provient des considérations précédentes).

Le point-clé pour établir l'existence d'un p-Sylow est le lemme suivant.

1.4. Lemme. *Soit G un groupe fini avec $|G| = n = p^m q$, $m \geqslant 0$, $p \nmid q$, et soit H un sous-groupe de G. Supposons que G admette un p-Sylow S. Alors, il existe $g \in G$ tel que $gSg^{-1} \cap H$ soit un p-Sylow de H.*

Démonstration. Soit $X = G/S$. Alors, G opère par translation sur G/S par l'exemple VII-1.5 (5). Cela induit alors une action de H sur G/S par restriction. Si $\overline{g} \in G/S$, on a alors

$$\mathrm{Stab}_H(\overline{g}) = \{h \in H \mid \overline{g} = \overline{hg}\} = \{h \in H \mid g^{-1}(hg) \in S\} = gSg^{-1} \cap H.$$

Son ordre est donc une puissance de p. En effet, gSg^{-1} est un p-Sylow de G par la remarque 1.3. Comme $gSg^{-1} \cap H$ est un sous-groupe de gSg^{-1}, c'est aussi un p-groupe (puisque l'ordre d'un sous-groupe divise l'ordre du groupe). Pour montrer le lemme, il suffit donc de vérifier qu'il existe $g \in G$ tel que $[H : gSg^{-1} \cap H]$ soit premier à p. D'après la proposition VII-1.12, on a

$$[H : gSg^{-1} \cap H] = \mathrm{card}(\mathcal{O}_{\overline{g}}).$$

Si toutes les orbites de G/S sous l'action de G avaient un cardinal divisible par p, alors $\mathrm{card}(G/S)$ serait aussi divisible par p, car $|G/S|$ est la somme des cardinaux des orbites distinctes. Or, S étant un p-Sylow du groupe G, on a $p \nmid [G : S]$. Ainsi, on obtient une contradiction, et l'on a fini. □

Nous allons maintenant montrer le résultat suivant.

1.5. Théorème (Sylow). *Soit G un groupe fini et soit p un nombre premier. Écrivons $|G| = p^m q$, avec $m \geqslant 0$ et $p \nmid q$. Alors :*

(1) *il existe des p-Sylow de G, et tout p-sous-groupe de G est contenu dans un p-Sylow ;*

(2) *le conjugué d'un p-Sylow de G est un p-Sylow de G, et tous les p-Sylow de G sont conjugués dans G. En particulier, si S est un p-Sylow de G, alors S est distingué dans G si, et seulement si, S est l'unique p-Sylow de G ;*

(3) *si N_p est le nombre de p-Sylow de G, on a $N_p \equiv 1\ [p]$ et $N_p \mid q$.*

Démonstration. Commençons par démontrer l'existence d'un p-Sylow.

D'après le théorème de Cayley, le groupe G s'identifie à un sous-groupe de $\mathfrak{S}_n$, où $n = |G|$. Maintenant, $\mathfrak{S}_n$ s'identifie lui-même à un sous-groupe de $\mathrm{GL}_n(\mathbb{F}_p)$ de la manière suivante. Si $\sigma \in \mathfrak{S}_n$, on pose $M_\sigma = (m_{ij})$, où

$$m_{\sigma(j)j} = \overline{1} \quad \text{et} \quad m_{ij} = \overline{0}, \qquad \text{si } i \neq \sigma(j).$$

On vérifie alors aisément que l'application

$$\begin{aligned} \mathfrak{S}_n &\longrightarrow \mathrm{GL}_n(\mathbb{F}_p) \\ \sigma &\longmapsto M_\sigma \end{aligned}$$

est un morphisme de groupes injectif. Ainsi, G est isomorphe à un sous-groupe G' de $\mathrm{GL}_n(\mathbb{F}_p)$. D'après le lemme précédent et l'exemple 1.2, G' admet un p-Sylow, et donc il en est de même de G.

Soit maintenant H un p-sous-groupe de G et soit S un p-Sylow de G. Par le lemme 1.4, il existe $g \in G$ tel que $gSg^{-1} \cap H$ soit un p-Sylow de H. Mais comme H est un p-groupe, on a donc $gSg^{-1} \cap H = H$. Ainsi, $H \subset gSg^{-1}$, qui est un p-Sylow de G d'après la remarque 1.3. Si maintenant H est un autre p-Sylow de G, on a alors $H = gSg^{-1}$ (puisque H et gSg^{-1} ont alors même ordre).

Si de plus S est le seul p-Sylow de G, on a donc $gSg^{-1} = S$ pour tout g dans G, et S est donc distingué dans G.

Réciproquement, si S est un p-Sylow de G qui est distingué dans G, alors on a l'inclusion $gSg^{-1} \subset S$. Comme ces deux groupes ont même ordre, on obtient donc $gSg^{-1} = S$. Comme tous les p-Sylow sont conjugués, on obtient que S est l'unique p-Sylow de G. Cela établit les points (1) et (2).

Montrons (3). Soit E l'ensemble des p-Sylow de G. D'après le point (2), l'action de G sur E par conjugaison est transitive. Autrement dit, si S est un p-Sylow de G, on a $\mathcal{O}_S = E$. Comme le cardinal d'une orbite divise l'ordre de G par la proposition VII-1.12, on a $|E| = N_p \mid n$. Remarquons maintenant que S opère aussi sur E, et que l'on a

$$N_p \equiv |E^S| \quad [p],$$

en vertu du théorème VII-2.6. Nous allons montrer que $E^S = \{S\}$ (S est bien entendu fixe sous l'action de S, puisque $sSs^{-1} = S$ pour tout $s \in S$). Soit T un p-Sylow, et supposons que $T \in E^S$. On a donc

$$sTs^{-1} = T, \qquad \text{pour tout } s \in S.$$

Soit G' le sous-groupe de G engendré par S et T. On a $S \subset G'$ et $T \subset G'$, et ce sont deux p-Sylow de G'. En effet, puisque G' est un sous-groupe de G, son ordre divise celui de G. On a donc

$$|G'| = p^\ell q', \quad \ell \in [\![0, m]\!], \ p \nmid q'.$$

Comme G' contient S, qui est d'ordre p^m, on a donc $\ell = m$, et les p-Sylow de G' sont donc d'ordre p^m. Puisque $gTg^{-1} \subset T$ pour tout $g \in S \cup T$, on a $xTx^{-1} \subset T$ pour tout $x \in G'$. Ainsi, T est un p-Sylow de G' qui est distingué dans G'. C'est donc le seul, d'où $T = S$. Ainsi, $\mathrm{card}(E^S) = 1$, et donc

$$N_p \equiv 1 \quad [p].$$

Puisque $N_p \mid p^m q$ et que N_p est donc premier à p (du fait de la congruence précédente), on en déduit que $N_p \mid q$. Cela achève la démonstration. □

Voici deux exemples d'application de ce théorème.

1.6. Exemple. Un groupe d'ordre 63 n'est pas simple.

En effet, soit G un tel groupe. Puisque $63 = 3^2 \cdot 7$, on a $N_7 \mid 9$, et donc nécessairement $N_7 = 1$, 3 ou 9. Mais on a aussi $N_7 \equiv 1$ [7], donc $N_7 = 1$. Ainsi, G n'admet qu'un 7-Sylow, qui est alors distingué dans G. Ainsi, G n'est pas simple.

1.7. Exemple. Un groupe d'ordre 33 est cyclique.

En effet, si G est d'ordre $33 = 3 \cdot 11$, on a donc $N_3 \mid 11$ et donc $N_3 = 1$ ou 11. Mais on a aussi $N_3 \equiv 1$ [3], donc $N_3 = 1$. Ainsi, un 3-Sylow H_1 est distingué dans G.

De même, on a $N_{11} \mid 3$ et $N_{11} \equiv 1$ [11], et donc $N_{11} = 1$. Par conséquent, un 11-Sylow H_2 est distingué dans G.

Comme 3 et 11 sont premiers entre eux, on obtient $H_1 \cap H_2 = \{1_G\}$. De plus, on a $|G| = |H_1| \cdot |H_2|$. Par le corollaire X-2.9, on a $G = H_1 \odot H_2$.

Comme 3 et 11 sont des nombres premiers, les groupes H_1 et H_2 sont cycliques. Soit h_1 un générateur de H_1 et soit h_2 un générateur de H_2. Alors, on a $o(h_1) = 3$ et $o(h_2) = 11$. Puisque $G = H_1 \odot H_2$, les éléments de H_1 commutent avec les éléments de H_2. On a alors $o(h_1 h_2) = 33$, d'après la proposition VI-4.17.

Le groupe G étant d'ordre 33, on en déduit que G est cyclique, engendré par $h_1 h_2$.

2. Actions primitives et critère de simplicité d'Iwasawa

Nous allons dans ce paragraphe introduire la notion d'action primitive d'un groupe, puis démontrer un critère de simplicité dû à Iwasawa.

On commence par définir la notion de relation d'équivalence G-stable.

2.1. Définition. Soit G un groupe agissant sur un ensemble E. Une relation d'équivalence (non vide) sur E est dite *G-stable* si pour tous x, $y \in E$ et tout $g \in G$, on a

$$x \sim y \quad \Longrightarrow \quad g \cdot x \sim g \cdot y.$$

En termes ensemblistes, une relation d'équivalence non vide $R \subset E \times E$ est donc G-stable si pour tout $(x, y) \in R$, et tout $g \in G$, on a $(g \cdot x, g \cdot y) \in R$.

2.2. Remarque. Si G agit sur E, alors l'application

$$\begin{aligned} G \times (E \times E) &\longrightarrow E \times E \\ \big(g, (x,y)\big) &\longmapsto g * (x,y) = (g{\cdot}x, g{\cdot}y) \end{aligned}$$

définit une action de G sur $E \times E$.

Une relation d'équivalence R non vide sera G-stable si, et seulement si, c'est une réunion d'orbites de $E \times E$ sous l'action de G.

En effet, une telle réunion d'orbites est bien G-stable. Inversement, si R est une relation d'équivalence non vide G-stable, alors pour tout $(x,y) \in R$, R contient la G-orbite de (x,y). On a ainsi

$$R = \bigcup_{(x,y)\in R} \mathcal{O}_{(x,y)}.$$

Attention ! Une réunion d'orbites n'est pas nécessairement une relation d'équivalence.

2.3. Exemples. Soit G un groupe agissant sur un ensemble E.

(1) Les parties $\Delta = \{(x,x) \mid x \in E\}$ et $E \times E$ sont clairement des relations d'équivalence G-stables. Elles correspondent respectivement à la relation d'égalité, et la relation d'équivalence grossière pour laquelle tous les éléments sont équivalents. Ces deux relations d'équivalence seront dites « triviales ».

(2) La relation d'équivalence « être dans la même G-orbite » est G-stable. En effet, pour tout $x \in E$, et tout $g \in G$, on a $g{\cdot}x \sim x$, car $g{\cdot}x$ est dans l'orbite de x par définition. En particulier, pour tous $x, y \in E$, et tout $g \in G$, on a

$$x \sim y \implies g{\cdot}x \sim x \sim y \sim g{\cdot}y.$$

Remarquons que si l'action de G sur E n'est ni triviale, ni transitive, cette relation d'équivalence est non triviale.

Nous allons maintenant nous intéresser aux actions d'un groupe G sur un ensemble E pour lesquelles les seules relations d'équivalence G-stables sur E sont triviales. D'après l'exemple précédent, l'action de G est nécessairement transitive dès qu'elle est non triviale.

2.4. Définition. Soit G un groupe agissant sur un ensemble E. On dit que l'action de G sur E est *primitive* (ou que G agit primitivement sur E) si elle est transitive, et si les seules relations d'équivalence G-stables sur E sont triviales.

En particulier, si $|E| \geqslant 2$, une action primitive de G est non triviale [1].

2.5. Exemples

(1) Si E est un singleton, toute action de G sur E est primitive.

(2) Soit E un ensemble possédant au moins trois éléments. L'action naturelle de $\mathfrak{S}(E)$ sur E est primitive.

Elle est bien entendu transitive. Posons $\Delta = \{(a,a) \mid E\}$. Alors, G agit transitivement sur $(E\times E)\setminus\Delta$. En effet, pour tous (a,b), $(c,d) \in E\times E$ tels que $a \neq b$ et $c \neq d$, il existe au moins une permutation $\sigma \in \mathfrak{S}(E)$ telle que $\sigma(a) = c$ et $\sigma(b) = d$. On peut prendre en effet

$$\sigma = \mathrm{Id}_E, \quad (b\ d), \quad (a\ c) \quad \text{ou} \quad (a\ c)(b\ d),$$

selon que $a = c$ et $b = d$, $a = c$ et $b \neq d$, $a \neq c$ et $b = d$, ou $a \neq c$ et $b \neq d$ respectivement.

Ainsi, $E \times E$ possède exactement deux orbites sous l'action de G, à savoir Δ et $(E \times E) \setminus \Delta$. Or, une relation d'équivalence G-stable R est réunion de telles orbites par la remarque 2.2, et contient nécessairement Δ (par réflexivité d'une relation d'équivalence). Par conséquent, on a $R = \Delta$ ou $R = E \times E$, et R est donc triviale.

(3) Soit $G = \langle(1\ 2\ 3\ 4)\rangle \subset \mathfrak{S}_4$ et $E = [\![1,4]\!]$. L'action naturelle de G sur E est transitive, mais pas primitive.

En effet, soit la relation $\sim$ sur E définie par : pour tous a, $b \in E$, $a \sim b$ si $a - b$ est pair.

Alors, $\sim$ est une relation d'équivalence non triviale, G-stable.

Nous allons maintenant décrire les relations d'équivalence G-stables sur E, lorsque l'action est transitive.

2.6. Théorème. *Soit G un groupe agissant transitivement sur un ensemble E, et soit $x_0 \in E$. Pour tout $x \in E$, on note g_x* **un** *élément de G tel que $x = g_x \cdot x_0$.*

(1) *Pour tout sous-groupe H de G contenant $\mathrm{Stab}_G(x_0)$, l'application*

$$\begin{aligned} E &\longrightarrow G/H \\ x &\longmapsto \overline{g}_x \end{aligned}$$

est bien définie. Soit $\sim_H$ la relation sur E définie comme suit : pour tous x, $y \in E$, $x \sim_H y$ si $\overline{g}_x = \overline{g}_y \in G/H$.

Alors, $\sim_H$ est une relation d'équivalence G-stable sur E bien définie.

1. En effet, l'action triviale sur un ensemble possédant au moins deux éléments n'est pas transitive.

(2) *Pour toute relation d'équivalence $\sim$ G-stable sur E, l'ensemble*

$$H_\sim = \{g \in G \mid g{\cdot}x_0 \sim x_0\}$$

est un sous-groupe de G contenant $\mathrm{Stab}_G(x_0)$.

De plus, les applications

$$\begin{array}{ccc} H & \longmapsto & \sim_H \\ H_\sim & \longleftarrow\!\shortmid & \sim \end{array}$$

établissent une correspondance bijective entre l'ensemble des sous-groupes de G contenant $\mathrm{Stab}_G(x_0)$ *et l'ensemble des relations d'équivalence sur E qui sont G-stables.*

Cette correspondance est (strictement) croissante. Autrement dit, pour tous sous-groupes H_1, H_2 de G contenant $\mathrm{Stab}_G(x_0)$, *on a $H_1 \subset H_2$ si, et seulement si, pour tous x, $y \in E$, on a*

$$x \sim_{H_1} y \quad \Longrightarrow \quad x \sim_{H_2} y.$$

Enfin, via cette correspondance, les sous-groupes $\mathrm{Stab}_G(x_0)$ *et G correspondent respectivement à la relation d'égalité et à la relation grossière.*

Démonstration. Commençons par démontrer le point (1), ainsi que la toute dernière assertion. Soit $x \in E$. Si g_1, $g_2 \in G$ vérifient $x = g_1{\cdot}x_0 = g_2{\cdot}x_0$, en faisant agir g_1^{-1} des deux côtés de l'égalité, on obtient $x_0 = g_1^{-1}g_2{\cdot}x_0$, soit encore $g_1^{-1}g_2 \in \mathrm{Stab}_G(x_0)$. Comme le sous-groupe H contient $\mathrm{Stab}_G(x_0)$, on a bien $g_1^{-1}g_2 \in H$, et donc $\overline{g}_1 = \overline{g}_2 \in G/H$. Ainsi, la classe à gauche $\overline{g}_x$ est indépendante du choix d'un élément $g_x \in G$ vérifiant $x = g_x{\cdot}x_0$, et l'application

$$\begin{array}{ccc} E & \longrightarrow & G/H \\ x & \longmapsto & \overline{g}_x \end{array}$$

est bien définie.

Il est alors clair que la relation $\sim_H$ est une relation d'équivalence sur E. Montrons maintenant qu'elle est G-stable. Soient x, $y \in E$ tels que $x \sim_H y$. Soient g_x, $g_y \in G$ tels que $x = g_x{\cdot}x_0$ et $y = g_y{\cdot}x_0$. Pour tout $g \in G$, on a alors

$$g{\cdot}x = g{\cdot}(g_x{\cdot}x_0) = gg_x{\cdot}x_0,$$

et de même, $g{\cdot}y = gg_y{\cdot}x_0$. Il s'agit donc de vérifier que $\overline{gg}_x = \overline{gg}_y$.

Par hypothèse, $\overline{g}_x = \overline{g}_y$, soit $g_x^{-1}g_y \in H$. Mais alors, on a

$$(gg_x)^{-1}(gg_y) = g_x^{-1}g_y \in H,$$

c'est-à-dire $\overline{gg}_x = \overline{gg}_y$, d'où la conclusion voulue.

Supposons maintenant que $H = \mathrm{Stab}_G(x_0)$. Dans ce cas, $x \sim_H y$ si, et seulement si, $g_y = g_x h$, avec $h \in \mathrm{Stab}_G(x_0)$. Mais alors, on obtient

$$y = (g_x h){\cdot}x_0 = g_x{\cdot}(h{\cdot}x_0) = g_x{\cdot}x_0 = x.$$

Ainsi, $\sim_H$ est la relation d'égalité dans ce cas.

Supposons enfin que $H = G$. Il est alors clair que $\sim_H$ est la relation grossière, puisqu'alors $g_x^{-1} g_y \in G$.

Montrons le point (2). Puisque $g \cdot x_0 = x_0$ pour tout $g \in \mathrm{Stab}_G(x_0)$ et que $\sim$ est réflexive, $H_\sim$ contient $\mathrm{Stab}_G(x_0)$, et contient donc 1_G en particulier. De plus, pour tous g_1, $g_2 \in H_\sim$, on a $g_1 \cdot x_0 \sim x_0$ et $g_2 \cdot x_0 \sim x_0$. Par G-stabilité, on obtient

$$(g_1 g_2) \cdot x_0 = g_1 \cdot (g_2 \cdot x_0) \sim g_1 \cdot x_0 \sim x_0.$$

De même, pour tout $g \in H_\sim$, par G-stabilité, on a

$$x_0 = g^{-1} \cdot (g \cdot x_0) \sim g^{-1} \cdot x_0,$$

d'où $g^{-1} \cdot x_0 \sim x_0$ par symétrie. Ainsi, $H_\sim$ est bien un sous-groupe de G contenant $\mathrm{Stab}_G(x_0)$.

Il reste à démontrer que les deux applications de l'énoncé sont réciproques l'une de l'autre, et croissantes (la stricte croissance provenant du caractère bijectif).

Montrons le caractère bijectif de la correspondance. Soit H un sous-groupe de G contenant $\mathrm{Stab}_G(x_0)$. Pour tout $g \in G$, on a

$$g \cdot x_0 \sim_H x_0 \iff \overline{g} = \overline{1}_G \in G/H \iff g \in H,$$

d'où l'égalité $H_{\sim_H} = H$. Enfin, soit $\sim$ une relation d'équivalence G-stable. Pour tous x, $y \in E$, on a

$$x \sim_{H_\sim} y \iff g_x^{-1} g_y \in H_\sim \iff g_x^{-1} g_y \cdot x_0 \sim x_0.$$

Par G-stabilité, cela équivaut à $g_y \cdot x_0 \sim g_x \cdot x_0$, soit $y \sim x$, ou encore $x \sim y$ par symétrie. Ainsi, $\sim_{H_\sim} = \sim$.

Enfin, soient H_1, H_2 deux sous-groupes de G contenant $\mathrm{Stab}_G(x_0)$, et tels que $H_1 \subset H_2$. Soient x, $y \in E$ tels que $x \sim_{H_1} y$. Alors, on a

$$g_x^{-1} g_y \in H_1 \subset H_2,$$

d'où $x \sim_{H_2} y$. Inversement, si $\sim_{H_1}$ et $\sim_{H_2}$ vérifient la condition du théorème, il est clair que $H_{\sim_1} \subset H_{\sim_2}$. Cela achève la démonstration. □

Rappelons maintenant la notion de sous-groupe maximal.

2.7. Définition. Soit G un groupe. On dit qu'un sous-groupe H de G est *maximal* dans G si $H \neq G$, et pour tout sous-groupe H', on a

$$H \subsetneq H' \implies H' = G.$$

Autrement dit, un sous-groupe maximal est un élément maximal de l'ensemble des sous-groupes propres de G pour la relation d'inclusion.

On peut démontrer qu'il existe des groupes sans sous-groupes maximaux (cf. exercice IX-15 **b**). Cependant, leur existence est assurée lorsque G est fini.

2.8. Lemme. *Soit G un groupe fini non trivial. Alors, tout sous-groupe strict de G est contenu dans un sous-groupe maximal de G.*

En particulier, G possède au moins un sous-groupe maximal.

Démonstration. Soit H_0 un sous-groupe strict de G. Alors, l'ensemble des sous-groupes stricts de G contenant H_0 est non vide (puisqu'il contient H_0). Choisissons un sous-groupe strict U de G contenant H_0 d'ordre maximal (un tel U existe puisque son ordre est majoré par celui de G).

Montrons que U est maximal. Par définition, $U \neq G$. Soit H un sous-groupe de G contenant strictement U. Alors, H contient H_0. Comme $|H| > |U|$, la définition de U montre que H ne peut être un sous-groupe strict de G. Par conséquent, $H = G$, et U est un sous-groupe maximal de G.

Le dernier point s'obtient en prenant $H_0 = \{1_G\}$. □

Le théorème précédent donne alors immédiatement une caractérisation des actions primitives d'un groupe (non nécessairement fini) en terme de stabilisateur.

2.9. Corollaire. *Soit G un groupe agissant sur E et soit $x_0 \in E$ fixé. Alors, l'action de G sur E est primitive si, et seulement si, G agit transitivement sur E et $\mathrm{Stab}_G(x_0)$ est un sous-groupe maximal de G.*

On peut maintenant énoncer le théorème d'Iwasawa.

2.10. Théorème (Iwasawa). *Soit G agissant sur E et soit $x_0 \in E$ fixé. On suppose que les conditions suivantes sont vérifiées :*

(1) *le groupe G agit primitivement sur E ;*

(2) *il existe un sous-groupe A distingué dans $\mathrm{Stab}_G(x_0)$ tel que*

$$G = \langle gag^{-1},\ g \in G,\ a \in A \rangle.$$

Alors, pour tout sous-groupe distingué N de G, soit N agit trivialement sur E, soit G/N est isomorphe à un quotient de A.

Démonstration. Gardons les hypothèses et les notations de l'énoncé, et posons $H = \mathrm{Stab}_G(x_0)$. Soit N un sous-groupe distingué de G.

Supposons tout d'abord que $N \subset H$. Alors, pour tout $\nu \in N$ et tout $g \in G$, on a $(g^{-1}\nu g)\cdot x_0 = x_0$, puisque N est distingué dans G. En faisant agir g

des deux côtés de l'égalité, on obtient

$$\nu\cdot(g\cdot x_0) = g\cdot x_0, \quad \text{pour tout } g \in G \text{ et tout } \nu \in N.$$

Comme G agit transitivement sur E, cela se récrit

$$\nu\cdot x = x, \quad \text{pour tout } x \in E \text{ et tout } \nu \in N.$$

Autrement dit, N agit trivialement sur E.

Supposons maintenant que N ne soit pas contenu dans H. Alors, le sous-groupe $\langle N, H\rangle$ contient strictement H. Puisque l'action de G sur E est primitive, H est un sous-groupe maximal de G par le corollaire précédent, et l'on obtient que $\langle N, H\rangle = G$. Par le théorème IX-2.5, on a alors $G = NH$, puisque le sous-groupe N est distingué dans G.

Or, pour tout $\nu \in N$, pour tout $h \in H$ et pour tout $a \in A$, on a

$$(\nu h)a(\nu h)^{-1} = \nu(hah^{-1})\nu^{-1} \in \langle N, A\rangle,$$

car A est distingué dans H. Comme G est engendré par les conjugués des éléments de A, on obtient $G = \langle N, A\rangle$. D'après le théorème IX-2.5, on a alors $G = NA$, et un isomorphisme de groupes

$$G/N = NA/N \simeq A/(N \cap A).$$

Cela achève la démonstration. □

Nous allons maintenant utiliser ce théorème pour donner des critères de simplicité. On a tout d'abord besoin d'une définition.

2.11. Définition. Soit G un groupe. Le *groupe dérivé*[2] de G est le sous-groupe $[G, G]$ engendré par les éléments

$$[x, y] = xyx^{-1}y^{-1}, \quad x, y \in G.$$

Un élément de G de la forme $[x, y]$ est appelé un *commutateur*.

Le lemme suivant donne une propriété extrêmement importante du groupe dérivé.

2.12. Lemme. *Soit G un groupe, et soit H sous-groupe distingué dans G tel que G/H soit abélien. Alors, H contient $[G, G]$.*

Démonstration. Soit H un sous-groupe distingué dans G tel que G/H soit abélien. Alors, pour tous g_1, $g_2 \in G$, on a

$$\overline{g_1 g_2 g_1^{-1} g_2^{-1}} = \overline{g}_1\overline{g}_2\overline{g}_1^{-1}\overline{g}_2^{-1} = \overline{g}_1\overline{g}_1^{-1}\overline{g}_2\overline{g}_2^{-1} = \overline{1}_G,$$

2. Ce sous-groupe a déjà été introduit dans l'exercice IX-17.

et donc $[g_1, g_2] \in H$ pour tous $g_1, g_2 \in G$. Par conséquent, $[G, G] \subset H$, d'où le résultat voulu. □

2.13. Remarque. Le lecteur trouvera d'autres propriétés du sous-groupe dérivé dans l'exercice IX-17, ainsi que dans le lemme 2.16, et une caractérisation des groupes résolubles via leurs groupes dérivés successifs dans l'exercice IX-23.

On peut maintenant tirer une première conséquence du théorème d'Iwasawa.

2.14. Corollaire. *Soit G agissant sur un ensemble E et soit $x_0 \in E$. On suppose que les conditions suivantes sont vérifiées :*

(1) *on a $G = [G, G]$;*

(2) *le groupe G agit primitivement sur E ;*

(3) *il existe un sous-groupe A distingué dans $\mathrm{Stab}_G(x_0)$, résoluble, tel que*

$$G = \langle gag^{-1},\ g \in G,\ a \in A \rangle.$$

Alors, soit tout sous-groupe propre distingué dans G agit trivialement sur E, soit G est simple.

Démonstration. Supposons que G ne soit pas simple, et soit N un sous-groupe propre distingué dans G. Supposons que N n'agisse pas trivialement sur E. Par le théorème d'Iwasawa, G/N est isomorphe à un quotient de A. Comme le groupe A est résoluble, tout quotient de A est résoluble par la proposition IX-3.21 (2), et donc G/N est résoluble. Comme $N \neq G$, G/N est un groupe résoluble non réduit au neutre. Considérons donc une suite de composition $(\overline{G}_0, \ldots, \overline{G}_r)$ de G/N, où chaque quotient successif est abélien. Quitte à supprimer les répétitions, on peut toujours supposer que $\overline{G}_1$ est un sous-groupe strict de $\overline{G}_0$. Remarquons que $r \geqslant 1$, car G/N est non trivial.

Si $r = 1$, $\overline{G}_0 = G/N$ est alors abélien (puisque $\overline{G}_r = \{\overline{1}_G\}$). D'après le lemme 2.12, on a $G = [G, G] \subset N$, d'où $N = G$, ce qui contredit le choix du sous-groupe N.

Si $r \geqslant 2$, le sous-groupe $\overline{G}_1$ est de la forme L/N, où L est un sous-groupe distingué de G contenant N, par la proposition IX-1.6. Remarquons que l'on a $L \neq G$ puisque $\overline{G}_1 \neq \overline{G}_0$. Par le théorème IX-2.5, on a un isomorphisme

$$\overline{G}_0/\overline{G}_1 = (G/N)/(L/N) \simeq G/L.$$

Ainsi, le groupe quotient G/L est abélien. Comme précédemment, on obtient la contradiction $L = G$.

Par conséquent, N agit trivialement sur E, ce qui achève la démonstration. □

On obtient alors le critère suivant.

2.15. Corollaire (Critère d'Iwasawa). *Soit G un groupe agissant sur un ensemble E, soit $\varphi : G \longrightarrow \mathfrak{S}(E)$ le morphisme de groupes correspondant et soit $x_0 \in E$ fixé. On suppose que les conditions suivantes sont vérifiées :*

(1) *on a $G = [G, G]$;*

(2) *le groupe G agit primitivement sur E ;*

(3) *il existe un sous-groupe A distingué dans $\mathrm{Stab}_G(x_0)$, résoluble, tel que*

$$G = \langle gag^{-1},\ g \in G,\ a \in A \rangle.$$

Alors, le groupe quotient $G/\mathrm{Ker}(\varphi)$ est simple.

Démonstration. Gardons les notations de l'énoncé. Si K est un sous-groupe de G, on note $\overline{K} = K/\mathrm{Ker}(\varphi) \cap K$. Nous allons montrer que $\overline{G}$ et $\overline{A}$ vérifient les hypothèses du corollaire précédent.

Pour tous g_1, $g_2 \in G$, on a $[\overline{g}_1, \overline{g}_2] = \overline{[g_1, g_2]}$, et l'on en déduit facilement que

$$[\overline{G}, \overline{G}] = \overline{[G, G]}.$$

Par hypothèse sur G, on obtient $[\overline{G}, \overline{G}] = \overline{G}$. Le lemme IX-2.2 montre que le morphisme

$$\overline{\varphi} : \overline{G} \longrightarrow \mathfrak{S}(E)$$

est injectif. Ce morphisme correspond à une action fidèle de $\overline{G}$ sur E. Comme $\overline{\varphi}(\overline{g}) = \varphi(g)$ pour tout $g \in G$, cette action est donnée par

$$\begin{aligned} \overline{G} \times E &\longrightarrow E \\ (\overline{g}, x) &\longmapsto \overline{g}\cdot x = g\cdot x. \end{aligned}$$

Puisque G agit transitivement sur E, $\overline{G}$ agit aussi transitivement sur E.

Si $\sim$ est une relation d'équivalence $\overline{G}$-stable sur E, elle est aussi G-stable, puisque par définition, $g\cdot x = \overline{g}\cdot x$, pour tout $x \in E$ et tout $g \in G$. Par hypothèse, $\sim$ est triviale. Ainsi, l'action de $\overline{G}$ de E est bien primitive.

Enfin, A étant résoluble, $\overline{A}$ est résoluble par la proposition IX-3.21 (2). De plus, G étant engendré par les conjugués d'éléments de A, $\overline{G}$ est engendré par les conjugués des éléments de $\overline{A}$.

Puisque l'action de $\overline{G}$ est fidèle, le seul sous-groupe de $\overline{G}$ agissant trivialement sur E est le sous-groupe $\{\overline{1}_G\}$. En particulier, il n'existe pas de sous-groupe propre distingué dans $\overline{G}$ agissant trivialement sur E. Par le corollaire 2.14, $\overline{G}$ est simple. □

Nous allons maintenant donner quelques résultats qui nous permettront d'appliquer le critère ci-dessus plus facilement. Commençons par un petit lemme qui permet de calculer $[G, G]$ dans certains cas.

2.16. Lemme. *Soit G un groupe, et soit G^2 le sous-groupe de G engendré par les carrés des éléments de G. Alors :*

(1) *si H est un sous-groupe distingué de G, $[H, H]$ est distingué dans G. En particulier, $[G, G]$ est distingué dans G ;*

(2) *on a $[G, G] \subset G^2$. De plus, si le groupe G est engendré par des éléments d'ordre 2, on a l'égalité $[G, G] = G^2$.*

Démonstration

(1) Soit H un sous-groupe distingué de G.
Pour tout $g \in G$, et tous x, $y \in H$, on a

$$\begin{aligned} g[x,y]g^{-1} &= g(xyx^{-1}y^{-1})g^{-1} \\ &= (gxg^{-1})(gyg^{-1})(gx^{-1}g^{-1})(gy^{-1}g^{-1}) \\ &= (gxg^{-1})(gyg^{-1})(gxg^{-1})^{-1}(gyg^{-1})^{-1} \\ &= [gxg^{-1}, gyg^{-1}] \in [H,H], \end{aligned}$$

car H est distingué dans G. Cela démontre en particulier que $[H, H]$ est distingué dans G. Le reste est alors clair.

(2) Pour tous x, $y \in G$, on a

$$[x,y] = xxx^{-1}yx^{-1}yy^{-1}y^{-1} = x^2(x^{-1}y)^2(y^{-1})^2 \in G^2.$$

On en déduit donc que $[G, G] \subset G^2$.

Supposons maintenant que G soit engendré par une partie P ne contenant que des éléments d'ordre 2. Remarquons tout d'abord que pour tout $x \in P$, et tout $g' \in G$, on a

$$(xg')^2 = xg'xg' = xg'x^{-1}g' = [x, g']g'^2,$$

puisque $x = x^{-1}$ pour tout $x \in P$.

De plus, tout élément du groupe G est en fait un produit d'éléments de P. Soit $g \in G$. Écrivons $g = x_1 \cdots x_r$, avec $x_1, \ldots, x_r \in P$. Une utilisation de la relation précédente et une récurrence facile montrent alors que $g^2 \in [G, G]$. On a donc $G^2 \subset [G, G]$, d'où l'égalité demandée. □

2.17. Exemple. Soit K un corps et soit E un K-espace vectoriel de dimension finie $n \geqslant 2$. On suppose que $|K| \geqslant 4$ si $n = 2$. Alors, on a

$$[\mathrm{SL}(E), \mathrm{SL}(E)] = \mathrm{SL}(E).$$

Comme tout élément de $\mathrm{SL}(E)$ est un produit de transvections d'après

la proposition IV-5.8, il suffit de montrer que toute transvection est un élément de $[\mathrm{SL}(E), \mathrm{SL}(E)]$.

Rappelons que deux transvections sont conjuguées dans $\mathrm{GL}(E)$ par la proposition IV-5.11 (1). Comme $\mathrm{SL}(E)$ est distingué dans $\mathrm{GL}(E)$, le sous-groupe $[\mathrm{SL}(E), \mathrm{SL}(E)]$ est distingué dans $\mathrm{GL}(E)$ par le lemme 2.16. Ainsi, si $[\mathrm{SL}(E), \mathrm{SL}(E)]$ contient une transvection, il les contient toutes.

Il reste donc à montrer que $[\mathrm{SL}(E), \mathrm{SL}(E)]$ contient au moins une transvection.

On suppose que $|K| \geqslant 4$. On peut alors choisir un élément $\lambda \in K \backslash \{0, 1, -1\}$. Choisissons une base $\mathbf{e}$ de E. Soient u, $v \in \mathrm{SL}(E)$ les endomorphismes de E dont les matrices représentatives dans la base $\mathbf{e}$ sont respectivement

$$U = \begin{pmatrix} \mathrm{I}_{n-2} & & \\ & \lambda & \\ & & \lambda^{-1} \end{pmatrix}, \quad V = \begin{pmatrix} \mathrm{I}_{n-2} & & \\ & 1 & 1 \\ & 0 & 1 \end{pmatrix},$$

et soit $w = [u, v] \in [\mathrm{SL}(E), \mathrm{SL}(E)]$. Un simple calcul montre que

$$\mathrm{Mat}(w\,; \mathbf{e}) = UVU^{-1}V^{-1} = \begin{pmatrix} \mathrm{I}_{n-2} & & \\ & 1 & \lambda^2 - 1 \\ & 0 & 1 \end{pmatrix}.$$

Comme $\lambda^2 - 1 \neq 0$, cela montre que w est une transvection.

On suppose que $|K| = 2$ ou 3, et $n \geqslant 3$. Soient u, $v \in \mathrm{SL}(E)$ les endomorphismes de E dont les matrices représentatives dans la base $\mathbf{e}$ sont respectivement

$$U = \begin{pmatrix} \mathrm{I}_{n-3} & & \\ & \begin{matrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{matrix} & \end{pmatrix}, \quad V = \begin{pmatrix} \mathrm{I}_{n-3} & & \\ & \begin{matrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{matrix} & \end{pmatrix},$$

et soit $w = [u, v] \in [\mathrm{SL}(E), \mathrm{SL}(E)]$. Cette fois-ci, on a

$$\mathrm{Mat}(w\,; \mathbf{e}) = UVU^{-1}V^{-1} = \begin{pmatrix} \mathrm{I}_{n-3} & & \\ & \begin{matrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{matrix} & \end{pmatrix}.$$

Ainsi, $\mathrm{Ker}(W - \mathrm{I}_n)$ est un hyperplan, et un calcul élémentaire donne l'égalité $(W - \mathrm{I}_n)^2 = 0$. La proposition IV-5.2 montre alors que w est une transvection.

Nous allons maintenant donner une condition suffisante sur une action de groupe pour qu'elle soit primitive.

2.18. Définition. Soit G agissant sur un ensemble E possédant au moins deux éléments.

On dit que l'action de G sur E est 2-*transitive* si pour tous $x, y, x', y' \in E$ tels que $x \neq x'$ et $y \neq y'$, il existe $g \in G$ tel que $g{\cdot}x = y$ et $g{\cdot}x' = y'$.

Autrement dit, l'action de G sur E est 2-transitive si l'action correspondante de G sur $E \times E \setminus \Delta$ est transitive.

2.19. Exemple. Si E est un ensemble possédant au moins trois éléments, l'exemple 2.5 (2) montre que l'action de $\mathfrak{S}(E)$ sur E est une action 2-transitive.

L'intérêt des actions 2-transitives pour le problème qui nous préoccupe est donné par la proposition suivante.

2.20. Proposition. *Soit G agissant sur un ensemble E non réduit à un singleton. Si cette action est 2-transitive, alors elle est primitive.*

Démonstration. L'action de G sur E étant 2-transitive, elle est transitive. En effet, pour tous $x, y \in E$, $x \neq y$, par 2-transitivité, il existe $g \in G$ tel que $g{\cdot}x = y$ et $g{\cdot}y = x$. En particulier, x et y sont dans la même orbite sous l'action de G. Soit $\sim$ une relation d'équivalence G-stable qui ne soit pas la relation d'égalité. On va montrer que $\sim$ est la relation grossière, i.e. que tous les éléments de E sont équivalents. Puisque $\sim$ n'est pas la relation d'égalité, il existe au moins un élément $x \in E$ dont la classe d'équivalence ne soit pas réduite à $\{x\}$. Nous allons démontrer que tout élément de E est alors équivalent à x. Soit $y \in E$. Par choix de x, il existe $z \in E, x \neq z$ tel que $x \sim z$. Si $z = y$, on a $x \sim y$. Si $z \neq y$, par 2-transitivité, il existe $g \in G$ tel que $g{\cdot}x = z$ et $g{\cdot}z = y$. Comme $x \sim z$, par G-stabilité, on obtient $z \sim y$, d'où $x \sim z \sim y$. Cela achève la démonstration. □

Nous allons maintenant appliquer le critère d'Iwasawa pour démontrer qu'une famille de groupes issus de l'algèbre linéaire est simple. Le lecteur verra d'autres applications des actions primitives à la simplicité dans les exercices.

2.21. Définition. Soit K un corps, et soit E un K-espace vectoriel de dimension finie. Le *groupe projectif spécial linéaire* est le groupe

$$\mathrm{PSL}(E) = \mathrm{SL}(E)/Z\big(\mathrm{SL}(E)\big).$$

On a alors le théorème suivant.

2.22. Théorème. *Soit K un corps, et soit E un K-espace vectoriel de dimension finie. Si $\dim_K(E) \geqslant 3$, ou si $\dim_K(E) = 2$ et $|K| \geqslant 4$, le groupe projectif spécial linéaire $\mathrm{PSL}(E)$ est simple.*

Démonstration. Soit X l'ensemble des droites de E. On a alors une action naturelle

$$\begin{aligned} \mathrm{SL}(E) \times X &\longrightarrow X \\ (u, D) &\longmapsto u(D). \end{aligned}$$

Cette action est 2-transitive.
En effet, soient D_1, D_2, Δ_1, $\Delta_2 \in X$ des droites de E, avec $D_1 \neq D_2$ et $\Delta_1 \neq \Delta_2$, et soient e_1, e_2, f_1, $f_2 \in E \setminus \{0\}$ des vecteurs directeurs respectifs. Comme $D_1 \neq D_2$, (e_1, e_2) est une famille libre, que l'on peut donc compléter en une base $\mathbf{e} = (e_1, \ldots, e_n)$ de E. De même, on peut compléter (f_1, f_2) en une base $\mathbf{f} = (f_1, \ldots, f_n)$ de E. Soit $u_0 \in \mathrm{GL}(E)$ l'unique endomorphisme de E tel que

$$u_0(e_i) = f_i, \quad \text{pour tout } i \in [\![1, n]\!].$$

Comme u_0 envoie une base sur une base, c'est un automorphisme de E. Soit $\lambda = \det(u_0)^{-1} \in K^\times$, et soit $u \in \mathrm{GL}(E)$ tel que

$$u(e_1) = \lambda f_1 \quad \text{et} \quad u(e_i) = f_i, \quad \text{pour tout } i \in [\![2, n]\!].$$

Alors, $\det(u) = \lambda \det(u_0) = 1$. Ainsi, on a $u \in \mathrm{SL}(E)$, et comme $u(e_1) = \lambda f_1$ et $u(e_2) = f_2$, on a bien $u(D_1) = \Delta_1$ et $u(D_2) = \Delta_2$.

Par la proposition 2.20, l'action de $\mathrm{SL}(E)$ sur X est primitive.

Si maintenant $u \in \mathrm{SL}(E)$ agit trivialement sur X, alors il stabilise toute droite de E. D'après l'exemple VI-2.12, u est alors une homothétie de rapport $\lambda \in K^\times$. Mais alors, comme $\det(u) = 1$, on obtient $\lambda^n = 1$, et finalement $u \in Z\big(\mathrm{SL}(E)\big)$, toujours par le même exemple. Ainsi, le noyau de l'action de G sur E est $Z\big(\mathrm{SL}(E)\big)$.

Soit D une droite de E quelconque fixée. Le stabilisateur H de D est donc

$$H = \{u \in \mathrm{SL}(E) \mid u(D) = D\}.$$

Soit $A = \{u \in \mathrm{SL}(E) \mid \mathrm{Im}(u - \mathrm{Id}_E) \subset D,\ u_{|D} = \mathrm{Id}_D\}$. Alors, A est un sous-groupe distingué de H.

En effet, fixons une base $\mathbf{e} = (e_1, \ldots, e_n)$ de E telle que $D = Ke_1$. Alors, on a

$$H = \Big\{u \in \mathrm{SL}(E) \mid \mathrm{Mat}(u\,; \mathbf{e}) = \begin{pmatrix} \det(M)^{-1} & * \\ 0 & M \end{pmatrix},\ M \in \mathrm{GL}_{n-1}(K)\Big\},$$

tandis que

$$A = \Big\{u \in \mathrm{SL}(E) \mid \mathrm{Mat}(u\,; \mathbf{e}) = \begin{pmatrix} 1 & * \\ 0 & \mathrm{I}_{n-1} \end{pmatrix}\Big\}.$$

Le calcul matriciel par blocs montre que A est un sous-groupe distingué

de H. On voit qu'il est abélien (donc résoluble) par calcul direct, ou en constatant que toute matrice comme ci-dessus est un produit de matrices de transvections $T_{1j}(\lambda)$, $\lambda \in K^\times$, $j \in [\![2, n]\!]$, et que toutes ces matrices commutent entre elles.

On remarque d'ailleurs que A contient toute les transvections t_λ dont la matrice représentative dans la base $\mathbf{e}$ est de la forme $T_{1,2}(\lambda)$, avec $\lambda \in K^\times$. Or, toute transvection est conjuguée à une telle transvection t_λ d'après la proposition IV-5.11.

Ainsi, le groupe G est engendré par les conjugués des éléments de A. Enfin, sous les hypothèses du théorème, on a $[\mathrm{SL}(E), \mathrm{SL}(E)] = \mathrm{SL}(E)$, d'après l'exemple 2.17. Par le critère d'Iwasawa, $\mathrm{PSL}(E)$ est simple. □

2.23. Remarque. On peut montrer que si $\dim_K(E) = 2$ et $|K| \leqslant 3$, le groupe $\mathrm{PSL}(E)$ n'est pas simple (cf exercice 20).

3. Exercices

E1. Montrer qu'un groupe d'ordre 255 n'est pas simple.

E2. Soit G un groupe d'ordre d'ordre $3553 = 11 \cdot 17 \cdot 19$.

a. Montrer que $N_{11} = N_{17} = N_{19} = 1$.

Soit H_p l'unique p-Sylow de G ($p = 11, 17, 19$).

b. Montrer que $G = H_{11} \odot H_{17} \odot H_{19}$.

c. En déduire que G est cyclique.

E3. Déterminer tous les groupes d'ordre 45 à isomorphisme près.

E4. Soient p, q deux nombres premiers distincts, $p < q$, et soit G un groupe d'ordre pq. On note H_ℓ un ℓ-Sylow de G ($\ell = p, q$).

a. Montrer que l'on a nécessairement $N_q = 1$, et que $N_p = 1$ ou q. Montrer que si $N_p = q$, alors nécessairement $p \mid q - 1$.

b. Montrer que $G \simeq \mathbb{Z}/q\mathbb{Z} \times_\rho \mathbb{Z}/p\mathbb{Z}$, pour un certain morphisme de groupes $\rho : \mathbb{Z}/p\mathbb{Z} \longrightarrow \mathrm{Aut}(\mathbb{Z}/q\mathbb{Z})$ (éventuellement trivial).

On rappelle que $\mathrm{Aut}(\mathbb{Z}/q\mathbb{Z}) \simeq (\mathbb{Z}/q\mathbb{Z})^\times$ et que ce groupe-ci est cyclique (cf. exercice VI-37 et corollaire 2.9).

c. On suppose que $p \nmid q - 1$. Montrer que $G = H_p \odot H_q$, puis que G est cyclique.

d. On suppose que $p \mid q - 1$.

(i) On suppose que ρ est trivial. Montrer que G est cyclique.

(ii) On suppose que ρ est non trivial. Montrer que tous les produits semi-directs $\mathbb{Z}/q\mathbb{Z} \times_\rho \mathbb{Z}/p\mathbb{Z}$ sont isomorphes.

Indication. Utiliser l'exercice X-8, ainsi que le théorème VI-6.5.

e. Quels sont les groupes d'ordre pq, à isomorphisme près ?

E5. Soit G un groupe d'ordre $255 = 3{\cdot}5{\cdot}17$.

a. Montrer que G possède un unique 17-Sylow H_{17}. Justifier l'existence de $y \in G$ tel que $H_{17} = \langle y \rangle$.

b. Quel est l'ordre du groupe G/H_{17} ? En utilisant les résultats de l'exercice précédent, en déduire l'existence de $x \in G$ tel que $G/H_{17} = \langle \overline{x} \rangle$.

c. Montrer que $15 \mid o(x)$. En déduire que $o(x) = 15$ ou 255.

d. On suppose que $o(x) = 255$. Quelle est la structure de G ?

e. On suppose que $o(x) = 15$.

(i) Montrer qu'il existe un entier $m \in [\![0, 16]\!]$ tel que $xyx^{-1} = y^m$, puis que $m^{15} \equiv 1 \mod 17$. En déduire que l'ordre de $[m]_{17}$ dans le groupe $(\mathbb{Z}/17\mathbb{Z})^\times$ divise 15, où $[m]_{17}$ désigne la classe de m modulo 17, puis que $m = 1$.

(ii) Que vaut $o(y)$? En déduire que G est cyclique.

f. Résumer le résultat obtenu.

E6. Soit $A = \langle a \rangle$ un groupe cyclique à trois éléments, soit $H = A \times A$ et soit $K = \mathfrak{S}_2$.

a. Pour tout $\pi \in K$, et tout $(a_1, a_2) \in H$, on pose

$$\rho_\pi\big((a_1, a_2)\big) = (a_{\pi^{-1}(1)}, a_{\pi^{-1}(2)}).$$

Montrer que $\rho_\pi \in \mathrm{Aut}(H)$ et que $\rho : K \longrightarrow \mathrm{Aut}(H)$ est un morphisme de groupes.

b. On pose $G = H \times_\rho K$.

(i) Quel est l'ordre de G ?

(ii) Justifier que G possède un unique 3-Sylow. Un 2-Sylow de G est-il unique ?

c. Déterminer tous les éléments d'ordre 2 de G, et en déduire le nombre de 2-Sylow de G.

d. On pose $\sigma = \big((a^2, a), \mathrm{Id}\big)$, $\tau = \big((a^2, a), (1\ 2)\big)$, $\gamma = \big((a, a), \mathrm{Id}\big)$.

Soit $D = \langle \sigma, \tau \rangle$. Montrer que $D \simeq D_3$ et que $G \simeq D_3 \times \mathbb{Z}/3\mathbb{Z}$.

E7. Soit p un diviseur premier de $|G|$, et soit X_p l'ensemble des p-sous-groupes de Sylow de G. On note N_p le nombre de p-Sylow de G.

a. Justifier que l'application

$$\begin{aligned} G \times X_p &\longrightarrow X_p \\ (g, S) &\longmapsto gSg^{-1} \end{aligned}$$

définit une action de G sur X_p, et que cette action est transitive.

Si S est un p-Sylow, on pose

$$N_G(S) = \{g \in G \mid gSg^{-1} = S\}.$$

On rappelle que $N_G(S)$ est un sous-groupe de G, et que S est distingué dans $N_G(S)$.

b. Montrer que $N_p = \dfrac{|G|}{|N_G(S)|}$, où S est un p-Sylow de G.

c. Soit $\varphi : G \longrightarrow \mathfrak{S}(X_p)$ le morphisme de groupes associé à l'action de G sur X_p définie en **a.** Donner une condition nécessaire et suffisante sur N_p pour que le morphisme φ soit non trivial.

d. Vérifier que l'on a

$$\mathrm{Ker}(\varphi) = \bigcap_{S \in X_p} N_G(S).$$

e. On suppose dans cette question que $|G| = pq$, avec $p \nmid q$. Montrer que le nombre d'éléments d'ordre p de G est $N_p(p-1)$.

Dans les exercices qui suivent, on pourra utiliser librement les résultats de l'exercice 7.

E8. Soit G un groupe simple. On suppose que $|G|$ est divisible par au moins deux nombres premiers distincts.

a. Soit p un nombre premier divisant $|G|$.

(i) Justifier l'existence d'un morphisme non trivial $\varphi_p : G \longrightarrow \mathfrak{S}_{N_p}$.

(ii) Montrer que φ_p est injectif.

(iii) Montrer que $\varepsilon \circ \varphi_p$ est trivial, où ε est le morphisme de signature. En déduire l'existence d'un morphisme injectif $G \hookrightarrow \mathfrak{A}_{N_p}$, puis que l'on a $|G|$ divise $\dfrac{N_p!}{2}$.

b. En utilisant les questions précédentes et l'exercice VIII-30, en déduire qu'un groupe simple d'ordre 60 est isomorphe à $\mathfrak{A}_5$.

E9. On se propose de démontrer qu'un groupe fini simple non abélien est d'ordre $\geqslant 60$. En particulier, $\mathfrak{A}_5$ est le plus petit groupe simple non abélien d'après l'exercice précédent.

On admet le théorème suivant[3], qui est une conséquence immédiate d'un résultat dû à Burnside, et qui sera démontré dans la dernière partie de cet ouvrage (cf. théorème XXXV-5.5).

Théorème. Soit G un groupe fini, dont l'ordre admet au plus deux diviseurs premiers. Alors, G n'est pas simple.

Soit G un groupe fini simple non abélien.

a. Montrer que $|G|$ possède au moins trois diviseurs premiers distincts.

b. En considérant le nombre de 3-Sylow et de 5-Sylow, montrer qu'un groupe d'ordre 30 n'est pas simple.

c. Conclure.

E10. On suppose que G est un groupe d'ordre 12, contenant un sous-groupe distingué isomorphe à $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

a. Donner la structure d'un 3-Sylow et d'un 2-Sylow. Montrer que le nombre N_3 de 3-Sylow est 1 ou 4.

b. Soit S un 3-Sylow et soit T un 2-Sylow. Montrer que $S \cap T = \{1_G\}$.

c. On suppose dans cette question que $N_3 = 1$. Montrer que $G = S \odot T$. En déduire que $G \simeq \mathbb{Z}/6\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

d. On suppose que $N_3 = 4$. On note $\varphi : G \longrightarrow \mathfrak{S}_4$ le morphisme induit par l'action de G sur l'ensemble de ses 3-Sylow.

(i) Montrer que pour tout 3-Sylow S, $N_G(S)$ est un 3-Sylow.

(ii) Montrer que deux 3-Sylow distincts s'intersectent trivialement.

(iii) En déduire que φ est injective et que G est isomorphe à un sous-groupe d'ordre 12 de $\mathfrak{S}_4$.

(iv) En déduire que $G \simeq \mathfrak{A}_4$.

e. Résumer le résultat obtenu.

E11. Soient p, q deux nombres premiers distincts, et soit G un groupe d'ordre pq^2.

a. On suppose que $p > q$. Montrer que $N_p = 1$ ou q^2. En déduire que, soit $N_p = 1$, soit $p = 3, q = 2$ et $N_3 = 4$.

b. On suppose que $q > p$. Montrer que $N_q = 1$.

3. On peut bien entendu s'en passer pour atteindre le but désiré, mais cela rallonge considérablement le nombre de cas à traiter.

c. On suppose que G est d'ordre 12 et que $N_3 = 4$. Montrer que $N_2 = 1$.

Indication. Compter les éléments.

d. En déduire qu'un groupe d'ordre pq^2 n'est pas simple.

E12. Soit p un nombre premier. On rappelle que pour tout $n \geqslant 1$, on a un isomorphisme de groupes

$$\mathrm{GL}_n(\mathbb{F}_p) \simeq \mathrm{Aut}\big((\mathbb{Z}/p\mathbb{Z})^n\big),$$

qui à une matrice inversible M fait correspondre le morphisme de multiplication à gauche par M.

a. Soit G le sous-groupe de $\mathrm{GL}_3(\mathbb{F}_p)$ constitué des matrices triangulaires supérieures possédant des $\overline{1}$ sur la diagonale. Montrer que

$$G \simeq (\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}) \times_\rho \mathbb{Z}/p\mathbb{Z},$$

où $\rho : \mathbb{Z}/p\mathbb{Z} \longrightarrow \mathrm{Aut}\big((\mathbb{Z}/p\mathbb{Z})^2\big)$ est un morphisme non trivial que l'on précisera.

b.

(i) Quel est l'ordre d'un p-Sylow de $\mathrm{GL}_2(\mathbb{F}_p)$?

(ii) En déduire que pour tous M, $M' \in \mathrm{GL}_2(\mathbb{F}_p)$ d'ordre p, il existe un entier $r \in \mathbb{Z}$ premier à p et $P \in \mathrm{GL}_2(\mathbb{F}_p)$ tels que $M' = PM^rP^{-1}$.

c. Déduire des questions précédentes qu'il existe un produit semi-direct non trivial

$$(\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z}) \times_\rho \mathbb{Z}/p\mathbb{Z},$$

unique à isomorphisme près.

d. Montrer que le centre de $(\mathbb{Z}/p\mathbb{Z}\times\mathbb{Z}/p\mathbb{Z})\times_\rho\mathbb{Z}/p\mathbb{Z}$ est isomorphe à $\mathbb{Z}/p\mathbb{Z}$.

E13. On rappelle qu'un sous-groupe G de $\mathfrak{S}_n$ est transitif s'il agit transitivement sur $[\![1, n]\!]$, et que dans ce cas $n \mid |G|$ (cf. exercice VIII-34).

Le but de cet exercice est déterminer tous les sous-groupes transitifs de $\mathfrak{S}_5$. On suppose connus les résultats des exercices X-12 et 7.

On note $\sigma = (1\ 2\ 3\ 4\ 5)$ et $\tau = (1\ 2\ 4\ 3)$, et l'on pose

$$S = \langle\sigma\rangle, \quad D = \langle\sigma, \tau^2\rangle, \quad N = \langle\sigma, \tau\rangle.$$

a. Soit G un sous-groupe de $\mathfrak{S}_5$. Montrer que si $G = \mathfrak{S}_5$, $\mathfrak{A}_5$ ou que G est conjugué à S, D ou N, alors G est transitif.

On note N_5 le nombre de 5-Sylow de G.

b. Montrer qu'un 5-Sylow de G est d'ordre 5, engendré par un 5-cycle. Montrer également que $N_5 = 1$ ou 6.

c. On suppose que $N_5 = 6$.

(i) En comptant les éléments d'ordre 5, montrer que G contient vingt-quatre 5-cycles.

(ii) En déduire que G contient tous les 5-cycles, puis que $\mathfrak{A}_5 \subset G$.

(iii) Conclure que $G = \mathfrak{A}_5$ ou $\mathfrak{S}_5$.

d. On suppose maintenant que $N_5 = 1$.

(i) Montrer que, quitte à remplacer G par un de ses conjugués, on peut supposer que $\sigma \in G$.

(ii) Soit $N_G(S) = \{g \in G \mid gSg^{-1} = S\}$. Montrer que $|N_G(S)| = 20$, et que $G \subset N_G(S)$.

(iii) Montrer que $N_G(S) = N$.

Indication. On a $N \simeq \mathrm{GLA}(\mathbb{F}_5)$, donc N est d'ordre 20.

(iv) Énumérer les sous-groupes de N, et en déduire que l'on a $G = S$, D ou N.

e. En déduire que les sous-groupes transitifs de $\mathfrak{S}_5$ sont $\mathfrak{S}_5$, $\mathfrak{A}_5$ et les conjugués de S, D et N. En particulier, montrer qu'un tel sous-groupe est égal à $\mathfrak{S}_5$, $\mathfrak{A}_5$ ou isomorphe à $\mathbb{Z}/5\mathbb{Z}$, D_5 ou $\mathrm{GLA}(\mathbb{F}_5)$.

E14. Soit p un nombre premier.

a. Dénombrer le nombre de p-cycles dans $\mathfrak{S}_p$.

b. En déduire le nombre de p-Sylow de $\mathfrak{S}_p$.

c. Montrer le théorème de Wilson : pour tout nombre premier p, on a

$$(p-1)! \equiv -1 \quad [p].$$

E15. Soit G un groupe abélien fini. Montrer que G est le produit direct interne de ses sous-groupes de Sylow. Est-ce vrai si G n'est pas abélien ?

E16. Soit G un groupe fini, soit p un nombre premier et soit S un p-Sylow de G.

a. Justifier que $Z(G)S$ est un sous-groupe de G, et que $(Z(G)S)/Z(G)$ est un p-Sylow de $G/Z(G)$.

On suppose dorénavant que $G/Z(G)$ est un p-groupe.

b. Montrer que $Z(G)S = G$.

c. Montrer que S est l'unique p-Sylow de G.

d. Montrer que $S \cap Z(G)$ est l'unique p-Sylow de $Z(G)$.

e. Montrer que $Z(G) = \big(S \cap Z(G)\big) \odot A$, où A est un groupe abélien d'ordre premier à p. On pourra utiliser l'exercice précédent.

f. Montrer que $G = S \odot A$.

E17. Soit G un groupe d'ordre 12.

a. Montrer que G est un produit semi-direct.

b. Montrer qu'il y a cinq groupes d'ordre 12 à isomorphisme près, à savoir

$$\mathbb{Z}/12\mathbb{Z}, \quad \mathbb{Z}/6\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}, \quad D_6, \quad \mathfrak{A}_4 \quad \text{et} \quad \mathbb{Z}/3\mathbb{Z} \times_\rho \mathbb{Z}/4\mathbb{Z},$$

où $\rho : \mathbb{Z}/3\mathbb{Z} \longrightarrow \mathrm{Aut}(\mathbb{Z}/3\mathbb{Z})$ est l'unique morphisme tel que

$$\rho_{\bar{1}}(\tilde{x}) = -\tilde{x}, \qquad \text{pour tout } \tilde{x} \in \mathbb{Z}/3\mathbb{Z}.$$

On prendra bien soin de montrer que ces groupes sont deux à deux non isomorphes.

On pourra utiliser les exercices X-10, X-8 et VI-38.

c. Reconnaître les groupes suivants dans la liste précédente :

(1) $\mathfrak{S}_3 \times \mathbb{Z}/2\mathbb{Z}$;

(2) $G = \langle A, B\rangle \subset \mathrm{GL}_2(\mathbb{C})$, où $A = \begin{pmatrix} j & 0 \\ 0 & j^2 \end{pmatrix}$ et $B = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$.

E18. On se propose de donner une classification des groupes d'ordre 30 à isomorphisme près.

a. Calculer le nombre d'éléments d'ordre au plus 2 des groupes

$$\mathbb{Z}/30\mathbb{Z}, \ D_{15}, \ \mathbb{Z}/5\mathbb{Z} \times D_3 \text{ et } \mathbb{Z}/3\mathbb{Z} \times D_5,$$

et en déduire que ces groupes sont deux à deux non isomorphes.

Dans toute la suite, on suppose que G est un groupe d'ordre 30.

b. Soit $p = 2$, 3 ou 5. On note N_p le nombre de p-Sylow de G.

(i) Calculer les valeurs possibles de N_3 et N_5.

(ii) Justifier qu'un p-Sylow de G est isomorphe à $\mathbb{Z}/p\mathbb{Z}$.

(iii) Montrer que deux p-Sylow de G distincts s'intersectent trivialement.

(iv) En déduire que G possède $N_p(p-1)$ éléments d'ordre p.

c. En utilisant la question précédente, montrer par l'absurde que $N_3 = 1$ ou $N_5 = 1$.

d. Pour $p = 2, 3, 5$, on note H_p un p-Sylow de G. Justifier qu'au moins un des deux sous-groupes H_3 et H_5 est distingué dans G. En déduire que $H = \langle H_3, H_5\rangle$ est un sous-groupe cyclique d'ordre 15.

e. Montrer que $G = H \rtimes H_2 \simeq H \times_\rho H_2$, où $\rho : H_2 \longrightarrow \mathrm{Aut}(H)$ est un morphisme de groupes.

f. On rappelle que si C est un groupe cyclique d'ordre n, alors on a un isomorphisme de groupes $\mathrm{Aut}(C) \simeq (\mathbb{Z}/n\mathbb{Z})^\times$ (cf. exercice VI-37).

(i) Lister les éléments d'ordre au plus 2 de $(\mathbb{Z}/15\mathbb{Z})^\times$.

(ii) En déduire qu'il y a au plus quatre groupes d'ordre 30 à isomorphisme près.

Indication. Si $\rho : H_2 \longrightarrow \mathrm{Aut}(H)$ est un morphisme de groupes, et si h est l'unique élément non trivial de H_2, que vaut $\rho(h)^2$?

g. Donner la liste de tous les groupes d'ordre 30, à isomorphisme près.

E19. Le but de cet exercice est de classer les groupes d'ordre 18 à isomorphisme près. On pourra utiliser les exercices X-8 et VI-38, ainsi que l'exercice VII-10.

Dans toute la suite, G désignera un groupe d'ordre 18.

a. On suppose que G est non commutatif, montrer que $G \simeq \mathbb{Z}/9\mathbb{Z} \times_\rho \mathbb{Z}/2\mathbb{Z}$ ou $G \simeq (\mathbb{Z}/3\mathbb{Z})^2 \times_\varphi \mathbb{Z}/2\mathbb{Z}$.

b. On suppose que $G \simeq \mathbb{Z}/9\mathbb{Z} \times_\rho \mathbb{Z}/2\mathbb{Z}$, où $\rho : \mathbb{Z}/2\mathbb{Z} \longrightarrow \mathrm{Aut}(\mathbb{Z}/9\mathbb{Z})$.

(i) Si ρ est trivial, montrer que $G \simeq \mathbb{Z}/18\mathbb{Z}$.

(ii) Montrer que $-\overline{1}$ est l'unique élément d'ordre 2 de $(\mathbb{Z}/9\mathbb{Z})^\times$.

(iii) En déduire qu'il y a un unique morphisme non trivial

$$\rho_0 : \mathbb{Z}/2\mathbb{Z} \longrightarrow \mathrm{Aut}(\mathbb{Z}/9\mathbb{Z})$$

que l'on explicitera, et que $G \simeq G_0 = \mathbb{Z}/9\mathbb{Z} \times_{\rho_0} \mathbb{Z}/2\mathbb{Z}$.

c. On suppose que $G \simeq (\mathbb{Z}/3\mathbb{Z})^2 \times_\varphi \mathbb{Z}/2\mathbb{Z}$.

(i) Si φ est trivial, montrer que $G \simeq \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}$.

(ii) Soit $M \in \mathrm{GL}_2(\mathbb{F}_3)$ un élément d'ordre 2. Montrer que $M = -\mathrm{I}_2$, ou que M est conjuguée à $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

(iii) En déduire que $G \simeq G_\varepsilon = (\mathbb{Z}/3\mathbb{Z})^2 \times_{\varphi_\varepsilon} \mathbb{Z}/2\mathbb{Z}$, où φ_ε est défini par

$$(\varphi_\varepsilon)_{\overline{1}} = \big((\overline{x}, \overline{y})\big) = (\overline{\varepsilon x}, -\overline{y}), \quad \text{pour tout } (\overline{x}, \overline{y}) \in (\mathbb{Z}/3\mathbb{Z})^2, \text{ avec } \varepsilon = \pm 1.$$

d.

(i) Montrer que G_0 possède deux éléments d'ordre 3.

(ii) Justifier que $(\mathbb{Z}/3\mathbb{Z})^2$ a huit éléments d'ordre 3.

(iii) Montrer que dans le groupe G_ε, pour tous $(\overline{x},\overline{y}) \in (\mathbb{Z}/3\mathbb{Z})^2$, on a

$$\big((\overline{x},\overline{y}),\overline{0}\big)^3 = \big((\overline{0},\overline{0}),\overline{0}\big).$$

En déduire que G_ε possède au moins huit éléments d'ordre 3.

(iv) Montrer que dans G_{-1}, on a

$$\big((\overline{x},\overline{y}),\overline{1}\big)^2 = \big((\overline{0},\overline{0}),\overline{0}\big), \quad \text{pour tous } (\overline{x},\overline{y}) \in (\mathbb{Z}/3\mathbb{Z})^2.$$

Quels sont les ordres possibles des éléments de G_{-1} ?

(v) Montrer que dans le groupe G_1, $\big((\overline{1},\overline{1}),\overline{1}\big)$ est d'ordre 6.

e. Montrer **très soigneusement** que $\mathbb{Z}/18\mathbb{Z}$, $\mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z}$, G_0, G_1 et G_{-1} sont deux à deux non isomorphes.

f. Combien il y a-t-il de classes d'isomorphisme de groupes d'ordre 18 ?

g. Identifier les groupes D_9 et $\mathfrak{S}_3 \times \mathbb{Z}/3\mathbb{Z}$ dans la liste obtenue.

Les exercices suivants nécessitent d'avoir étudié le paragraphe §2 de ce chapitre.

E20. Soit E un K-espace vectoriel de dimension 2, et soit $G = \mathrm{PSL}(E)$. On suppose[4] que $|K| = q$.

a. Vérifier rapidement que l'action de $\mathrm{SL}(E)$ sur les droites de E induit un morphisme injectif

$$\mathrm{PSL}(E) \hookrightarrow \mathfrak{S}_{q+1}.$$

b. On suppose que $q = 2$. Montrer que $|G| = 6$. En déduire que $\mathrm{PSL}(E)$ est isomorphe à $\mathfrak{S}_3$ et que $\mathrm{PSL}(E)$ n'est pas simple.

c. On suppose que $q = 3$. Montrer que $|G| = 12$. En déduire que $\mathrm{PSL}(E)$ est isomorphe à $\mathfrak{A}_4$ et que $\mathrm{PSL}(E)$ n'est pas simple.

Dans les deux exercices suivants, on se propose de donner une autre démonstration du théorème VIII-5.3 et du corollaire VIII-5.7. On s'interdira donc d'utiliser ces résultats.

E21. Soit E un ensemble fini.

a. On suppose que $|E| \geqslant 4$. Montrer que l'action du groupe alterné $\mathfrak{A}(E)$ sur E est 2-transitive.

b. On suppose que $|E| \geqslant 5$.

4. Les corps finis seront étudiés dans un chapitre ultérieur.

(i) Montrer que les produits de deux transpositions à supports disjoints engendrent $\mathfrak{A}(E)$.

(ii) Montrer que $[\mathfrak{A}(E), \mathfrak{A}(E)] = \mathfrak{A}(E)^2$.

Indication. Si σ est un 3-cycle, que vaut σ^4?

(iii) En déduire que $[\mathfrak{A}(E), \mathfrak{A}(E)] = \mathfrak{A}(E)$.

(iv) Montrer que les produits de deux transpositions à supports disjoints sont conjugués dans $\mathfrak{A}(E)$.

(v) Ces résultats sont-ils encore vrais si $|E| = 4$?

c. On suppose que $|E| = 5$.

(i) Soit $a \in E$, et soit H le stabilisateur de a. Soit A le sous-groupe engendré par les produits de deux transpositions à supports disjoints, et dont les supports ne contiennent pas a. Montrer que A est abélien, et distingué dans H.

(ii) Déduire des questions précédentes que $\mathfrak{A}(E)$ est simple.

E22. Soit E un ensemble vérifiant $|E| \geqslant 5$, et soit X l'ensemble des parties P de E à deux éléments.

a. Montrer que l'application

$$\begin{aligned} \mathfrak{S}(E) \times X &\longrightarrow X \\ (\sigma, P) &\longmapsto \sigma(P) \end{aligned}$$

est une action de $\mathfrak{S}(E)$ sur X, et que cette action est fidèle et transitive.

b. Montrer que $X \times X$ possède exactement trois orbites $\mathcal{O}_0$, $\mathcal{O}_1$, $\mathcal{O}_2$ sous l'action de $\mathfrak{S}(E)$, à savoir

$$\mathcal{O}_i = \{(P_1, P_2) \in X \times X \mid |P_1 \cap P_2| = i\}, \quad i \in [\![0, 2]\!] .$$

c. En déduire que $\mathfrak{S}(E)$ agit primitivement sur X.

Indication. On montrera que si une relation d'équivalence $\mathfrak{S}(E)$-stable contient un élément de $\mathcal{O}_0$, elle contient un élément de $\mathcal{O}_1$, et réciproquement.

d. Soit $P = \{a, b\} \in X$. Montrer que $\mathrm{Stab}_{\mathfrak{S}(E)}(P) = \langle (a\ b),\ \mathfrak{S}(E \setminus P) \rangle$.

e. Montrer directement que $\mathrm{Stab}_{\mathfrak{S}(E)}(P)$ est un sous-groupe maximal du groupe $\mathfrak{S}(E)$, et retrouver le fait que $\mathfrak{S}(E)$ agit primitivement sur X.

f. Montrer que $[\mathfrak{S}(E), \mathfrak{S}(E)] = \mathfrak{A}(E)$.

g. En utilisant le théorème d'Iwasawa et les questions précédentes, en déduire que un sous-groupe distingué de $\mathfrak{S}(E)$ non réduit à $\{\mathrm{Id}_E\}$ contient $\mathfrak{A}(E)$.

h. En conclure que les seuls sous-groupes distingués de $\mathfrak{S}(E)$ sont

$$\{\mathrm{Id}_E\}, \quad \mathfrak{A}(E) \quad \text{et} \quad \mathfrak{S}(E).$$

i. On suppose que $|E| \geqslant 6$. Reprendre les questions précédentes en faisant agir cette fois le groupe $\mathfrak{A}(E)$, et en déduire que $\mathfrak{A}(E)$ est simple.

E23. Cet exercice nécessite de connaître la définition d'une suite de Jordan-Hölder.

Soit K un corps infini, et soit E un K-espace vectoriel de dimension finie $n \geqslant 2$. Montrer que $\mathrm{PSL}(E)$ admet une suite de Jordan-Hölder, et exhiber un sous-groupe de $\mathrm{PSL}(E)$ qui n'en admette pas.

E24. Soit K un corps, et soit E un K-espace vectoriel de dimension $n \geqslant 1$.

a. Vérifier que tout sous-groupe de $Z\big(\mathrm{SL}(E)\big)$ est distingué dans $\mathrm{SL}(E)$.

b. On suppose que $|K| \geqslant 4$ si $n = 2$. Soit H un sous-groupe distingué de $\mathrm{SL}(E)$. Pour alléger les notations, on pose $Z = Z\big(\mathrm{SL}(E)\big)$.

(i) Montrer que, soit H est contenu dans Z, soit $\mathrm{SL}(E) = HZ$.

(ii) Montrer que si $\mathrm{SL}(E) = HZ$, alors $\mathrm{SL}(E)/H$ est abélien. En déduire alors que $H = \mathrm{SL}(E)$ (on pourra utiliser l'exercice IX-17).

(iii) En déduire que les sous-groupes distingués de $\mathrm{SL}(E)$ sont $\mathrm{SL}(E)$ et les sous-groupes de $Z\big(\mathrm{SL}(E)\big)$.

c. Décrire tous les sous-groupes distingués de $\mathrm{SL}_n(\mathbb{R})$ et de $\mathrm{SL}_n(\mathbb{C})$.

d. Justifier que $\mathrm{SL}_2(\mathbb{F}_2) \simeq \mathfrak{S}_3$. En déduire que le résultat de **b** (iii) est encore vrai dans ce cas.

e. Montrer que le sous-groupe de $\mathrm{SL}_2(\mathbb{F}_3)$ engendré par les matrices

$$A = \begin{pmatrix} \overline{0} & \overline{1} \\ -\overline{1} & \overline{0} \end{pmatrix} \quad \text{et} \quad B = \begin{pmatrix} -\overline{1} & -\overline{1} \\ -\overline{1} & \overline{1} \end{pmatrix}$$

est distingué dans $\mathrm{SL}_2(\mathbb{F}_3)$, et isomorphe à Q_8.

Le résultat de **b** (iii) est-il encore vrai dans ce cas ?

E25. Soit K un corps, et soit E un K-espace vectoriel de dimension $n \geqslant 1$. On suppose connus les résultats de l'exercice 24.

a. Soit S un sous-groupe de $K^\times$. Vérifier que les sous-groupes $\det^{-1}(S)$ et $S\mathrm{Id}_E$ sont distingués dans $\mathrm{GL}(E)$. Montrer que ce sont les seuls sous-groupes distingués dans $\mathrm{GL}(E)$ si $n = 1$.

b. On suppose que $n \geqslant 2$, et que $|K| \geqslant 4$ si $n = 2$. Soit H un sous-groupe distingué de $\mathrm{GL}(E)$.

(i) Justifier que $H \cap \mathrm{SL}(E)$ est soit contenu dans le centre de $\mathrm{SL}(E)$, soit égal à $\mathrm{SL}(E)$.

(ii) On suppose que $H \cap \mathrm{SL}(E) = \mathrm{SL}(E)$. Justifier que $S = \det(H)$ est un sous-groupe de $K^\times$, et montrer alors que $H = \det^{-1}(S)$.

(iii) On suppose que $H \cap \mathrm{SL}(E)$ est contenu dans le centre de $\mathrm{SL}(E)$. Soit $u \in H$ et soit $t \in \mathrm{SL}(E)$ une transvection de E. Montrer qu'il existe $\lambda \in K^\times$ tel que $utu^{-1} = \lambda t$.
Justifier que utu^{-1} est une transvection de E. En déduire que $\lambda = 1$, puis que u commute avec toute transvection. Montrer alors que tout élément de H est une homothétie.

(iv) En déduire que les sous-groupes distingués de $\mathrm{GL}(E)$ sont les sous-groupes $S\mathrm{Id}_E$ et les sous-groupes $\det^{-1}(S)$, lorsque S parcourt l'ensemble des sous-groupes de $K^\times$.

c. Justifier que $\mathrm{GL}_2(\mathbb{F}_2) \simeq \mathfrak{S}_3$. En déduire que le résultat de **b** (iv) est encore vrai dans ce cas.

d. Montrer que le résultat de **b** (iv) n'est plus vrai lorsque l'on suppose que $n = 2$ et $|K| = 3$. On pourra considérer le sous-groupe de $\mathrm{SL}_2(\mathbb{F}_3)$ introduit dans la question **e.** de l'exercice XI-24.

E26. On souhaite montrer que $-I_2$ n'est pas un commutateur de $\mathrm{SL}_2(\mathbb{C})$.
On suppose au contraire que $-I_2 = [A, B]$, avec $A, B \in \mathrm{SL}_2(\mathbb{C})$.

a. Montrer que $B = -ABA^{-1}$, et calculer $\mathrm{tr}(B)$.

b. Justifier que l'endomorphisme de $\mathbb{R}^2$ représenté par B n'est pas une homothétie, puis montrer que B est semblable à $C = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$.

c. En déduire que l'on peut supposer $B = C$ sans perte de généralité. Calculer alors $[A, C]$ (ne pas oublier que $\det(A) = 1$ pour les calculs), et conclure.

d. On pose $U = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix}$.
Pour $\mu \in \mathbb{C}^\times$, calculer $[U, T_{1,2}(\mu)]$ et $[U, T_{2,1}(\mu)]$, et utiliser les calculs faits dans l'exemple XI-2.17 pour écrire $-I_2$ comme un produit de commutateurs de $\mathrm{SL}_2(\mathbb{C})$.

E27. On suppose connus les résultats de l'exercice VII-14. Soit G un groupe fini.

a. On suppose que $G \simeq (\mathbb{Z}/2\mathbb{Z})^m$, $m \geqslant 1$ ou bien que $G \simeq \mathbb{Z}/3\mathbb{Z}$. Montrer que l'action de $\mathrm{Aut}(G)$ sur $G \setminus \{1_G\}$ est 2-transitive.

On s'intéresse maintenant à la réciproque et l'on suppose que l'action de $\mathrm{Aut}(G)$ sur $G \setminus \{1_G\}$ est 2-transitive.

b. Justifier que $G \simeq (\mathbb{Z}/p\mathbb{Z})^m$, où p est premier et $m \geqslant 1$.

c. Montrer que si $m \geqslant 2$, alors $p = 2$.

Indication. Considérer G comme un $\mathbb{F}_p$-espace vectoriel, et justifier l'existence de deux éléments v, $w \in G$ linéairement indépendants. Si p est impair, utiliser la 2-transitivité de l'action de $\mathrm{Aut}(G)$ pour en déduire une contradiction en considérant les éléments $v, -v$ et w.

d. On suppose que $m = 1$, i.e. $G \simeq \mathbb{Z}/p\mathbb{Z}$. Montrer que le seul automorphisme de G fixant un élément non nul est Id_G. Grâce à la 2-transitivité de l'action de $\mathrm{Aut}(G)$, en déduire que G possède au plus deux éléments non nuls, puis que $G \simeq \mathbb{Z}/2\mathbb{Z}$ ou $\mathbb{Z}/3\mathbb{Z}$.

e. Résumer le résultat obtenu [5].

5. Le résultat de cet exercice est tiré de [21].

Chapitre XII

Thème et variations sur les groupes abéliens

1. Torsion dans un groupe abélien

1.1. Définition. Si G un groupe quelconque, on note $\mathrm{Tor}(G)$ l'ensemble des éléments d'ordre fini de G. On dit que G est un *groupe de torsion* si l'on a $G = \mathrm{Tor}(G)$ (c'est-à-dire si tout élément de G est d'ordre fini), et est un *groupe sans torsion* si $\mathrm{Tor}(G) = \{1_G\}$. Un groupe G qui n'est ni de torsion, ni sans torsion sera appelé un *groupe mixte*.

1.2. Exemples

(1) Le groupe $\mathbb{Z}$ est sans torsion.

(2) Tout groupe fini est de torsion.

(3) Le groupe $\mathbb{Q}/\mathbb{Z}$ est de torsion, puisque pour tout $x \in \mathbb{Q}$, il existe $m \geqslant 1$ tel que $mx \in \mathbb{Z}$. Ce groupe est infini, puisqu'il contient au moins un élément d'ordre n pour tout $n \geqslant 1$, à savoir $\overline{\frac{1}{n}}$.

(4) Le groupe $\mathbb{Z} \times \mathbb{Q}/\mathbb{Z}$ est mixte.

(5) Le groupe $\mathrm{GL}_2(\mathbb{C})$ est mixte.

En effet, $A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ est d'ordre infini puisque les coefficients de A^n sont strictement positifs pour tout $n \geqslant 1$. En particulier, $A^n \neq \mathrm{I}_2$ pour tout $n \geqslant 1$. Par ailleurs, $-\mathrm{I}_2$ est un élément d'ordre 2.

En général, l'ensemble $\mathrm{Tor}(G)$ n'est pas un sous-groupe de G, comme le montre l'exemple suivant.

1.3. Exemple. Soit $G = \mathrm{GL}_2(\mathbb{C})$.

Alors, les matrices $M = \begin{pmatrix} 0 & -1 \\ 1 & -1 \end{pmatrix}$ et $N = M^t$ sont d'ordre 3, mais le produit $MN = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}$ est d'ordre infini, d'après l'exemple précédent.

Ainsi, $\mathrm{Tor}(G)$ n'est pas un sous-groupe de G.

Tout se passe mieux en revanche dans le cas d'un groupe abélien.

1.4. Lemme. *Soit G un groupe abélien. Alors, $\mathrm{Tor}(G)$ est un sous-groupe de torsion de G, et $G/\mathrm{Tor}(G)$ est un groupe sans torsion.*

Démonstration. Montrons que $\mathrm{Tor}(G)$ est un sous-groupe de torsion de G. Puisque $1_G \in \mathrm{Tor}(G)$, $\mathrm{Tor}(G)$ est non vide. Soient x, $y \in \mathrm{Tor}(G)$. Alors, il existe deux entiers n, $m \geqslant 1$ tels que $x^m = y^n = 1_G$. Mais alors, on a

$$(xy^{-1})^{mn} = x^{mn}y^{-mn} = (x^m)^n(y^n)^{-m} = 1_G,$$

la première égalité découlant du fait que G est abélien. Ainsi, $\mathrm{Tor}(G)$ est un sous-groupe de G. De plus, tout élément de $\mathrm{Tor}(G)$ est d'ordre fini par définition ; c'est donc un groupe de torsion.

Enfin, montrons que $G/\mathrm{Tor}(G)$ est sans torsion. Soit $\overline{x} \in G/\mathrm{Tor}(G)$ un élément d'ordre fini. Il existe donc $m \geqslant 1$ tel que $\overline{x}^m = \overline{x^m} = \overline{1}_G$. On a ainsi $x^m \in \mathrm{Tor}(G)$, et par suite il existe $n \geqslant 1$ tel que $(x^m)^n = 1_G$. Mais alors, x est d'ordre fini, et ainsi $x \in \mathrm{Tor}(G)$, d'où $\overline{x} = \overline{1}_G$. Par conséquent, le groupe $G/\mathrm{Tor}(G)$ est sans torsion. Cela achève la démonstration. □

1.5. Remarque. Soient G, G' deux groupes abéliens. Si $G \simeq G'$, on a alors un isomorphisme $\mathrm{Tor}(G) \simeq \mathrm{Tor}(G')$.

En effet, soit $\varphi : G \xrightarrow{\sim} G'$ un isomorphisme. Si $x \in \mathrm{Tor}(G)$, il existe $m \geqslant 1$ tel que $x^m = 1_G$. Mais alors, on a

$$\varphi(x)^m = \varphi(1_G) = 1_{G'},$$

et ainsi $\varphi(x) \in \mathrm{Tor}(G')$. Le morphisme φ se restreint donc en un morphisme de groupes $\mathrm{Tor}(G) \longrightarrow \mathrm{Tor}(G')$. De même, φ^{-1} se restreint en un morphisme $\mathrm{Tor}(G') \longrightarrow \mathrm{Tor}(G)$, et ces deux morphismes sont inverses l'un de l'autre, d'où le résultat.

1.6. Définition. Soit G un groupe abélien et soit p un nombre premier. On dit que G est un groupe *p-primaire* si G est de torsion et tout élément de G est d'ordre une puissance de p.

La *composante p-primaire* de G est le sous-ensemble

$$G_p = \{x \in G \mid \text{ il existe } m \geqslant 1 \text{ tel que } x^{p^m} = 1_G\}.$$

Autrement dit, c'est le sous-ensemble des éléments de G dont l'ordre est fini et égal à une puissance de p.

1.7. Proposition. *Soit G un groupe abélien et soit p un nombre premier. Alors, la composante p-primaire G_p de G est un sous-groupe p-primaire du groupe G.*

De plus :

(1) *si G_p est fini, G_p est un p-groupe ;*

(2) *si G est fini, G_p est l'unique p-Sylow de G, et l'on a*

$$G_p = \{x \in G \mid x^{p^r} = 1_G\},$$

où r est la plus grande puissance de p divisant l'ordre de G.

Démonstration. Comme $1_G \in G_p$, G_p est non vide. Soient x, $y \in G_p$. Alors, il existe $m, n \geqslant 1$ tels que $x^{p^m} = y^{p^n} = 1_G$ et ainsi

$$(xy^{-1})^{p^{m+n}} = (x^{p^m})^{p^n}(y^{p^n})^{-p^m} = 1_G.$$

Par conséquent, G_p est un sous-groupe de G, p-primaire par définition.

Supposons que G_p soit un groupe fini. Si l'ordre de G_p n'est pas une puissance de p, $|G_p|$ possède un diviseur premier $\ell \neq p$. Par le théorème de Cauchy, G_p contient un élément d'ordre ℓ, ce qui est absurde, car ℓ n'est pas une puissance de p. Ainsi, $|G_p|$ est une puissance de p, d'où (1).

Supposons maintenant que G soit fini, d'ordre $p^r q$, $p \nmid q$. Soit S un p-Sylow de G. On a donc $|S| = p^r$ et par le corollaire VI-4.15, on a

$$S \subset \{x \in G \mid x^{p^r} = 1_G\}.$$

De plus, S est un sous-groupe distingué de G, car G est abélien, et est donc l'unique p-Sylow de G par le théorème XI-1.5 (2). Puisque G_p est un p-groupe, il est contenu dans S, d'après le théorème XI-1.5 (1). Mais, on a $\{x \in G \mid x^{p^r} = 1_G\} \subset G_p$, puisque tout élément de l'ensemble de gauche est d'ordre divisant p^r, donc est d'ordre une puissance de p. Ainsi, on a finalement

$$G_p \subset S \subset \{x \in G \mid x^{p^r} = 1_G\} \subset G_p,$$

d'où les égalités

$$G_p = S = \{x \in G \mid x^{p^r} = 1_G\},$$

ce qui achève la démonstration. □

L'intérêt des composantes p-primaires est donné par le théorème suivant.

1.8. Théorème. *Soit G un groupe abélien. Alors, on a*

$$\mathrm{Tor}(G) = \bigoplus_p G_p \simeq \coprod_p G_p,$$

où p parcourt l'ensemble des nombres premiers.

En particulier, tout groupe abélien de torsion est somme directe interne de ses composantes p-primaires.

Démonstration. Soit G un groupe abélien. Soit $x \in \mathrm{Tor}(G)$. Écrivons

$$o(x) = p_1^{m_1} \cdots p_r^{m_r},$$

où $r \geqslant 0$, $p_1, \ldots, p_r$ sont des nombres premiers distincts, et $m_1, \ldots, m_r \geqslant 1$. Pour tout $i \in [\![1, r]\!]$, on pose

$$q_i = \frac{o(x)}{p_i^{m_i}}.$$

Les entiers $q_1, \ldots, q_r$ étant premiers entre eux dans leur ensemble, il existe des entiers $u_1, \ldots, u_r \in \mathbb{Z}$ tels que

$$u_1 q_1 + \cdots + u_r q_r = 1.$$

On remarque alors que, pour tout $i \in [\![1, r]\!]$, on a

$$(x^{u_i q_i})^{p_i^{m_i}} = x^{u_i o(x)} = 1_G,$$

et donc que $x^{u_i q_i} \in G_{p_i}$. Comme $x = x^1 = (x^{u_1 q_1}) \cdots (x^{u_r q_r})$, cela démontre que x s'écrit sous la forme

$$x = \prod_p x_p, \quad x_p \in G_p,$$

où les x_p sont presque tous égaux à 1_G.

Il reste à voir qu'une telle décomposition est unique. Supposons que l'on ait deux décompositions

$$x = \prod_p x_p = \prod_p x'_p.$$

Comme G est abélien, on en déduit donc que

$$\prod_p x'_p x_p^{-1} = 1_G.$$

Puisque G_p est un sous-groupe de G, on a $y_p = x'_p x_p^{-1} \in G_p$. Il s'agit donc de montrer que $y_p = 1_G$ pour tout p premier.

Par hypothèse, on a $y_p = 1_G$, sauf éventuellement pour un nombre fini de nombres premiers $p_1, \ldots, p_r$. On a donc

$$y_{p_1} \cdots y_{p_r} = 1_G.$$

Il suffit donc de démontrer par récurrence la propriété (H_n) suivante :

(H_n) Pour tous nombres premiers deux à deux distincts $p_1 \ldots, p_n$, et pour tous $z_1 \in G_{p_1}, \ldots, z_n \in G_{p_n}$, on a

$$z_1 \cdots z_n = 1_G \quad \Longrightarrow \quad z_1 = \cdots = z_n = 1_G.$$

La propriété (H_1) est clairement vérifiée. Supposons que (H_n) soit vraie pour un $n \geqslant 1$, et montrons que (H_{n+1}) est vraie. Pour tout $i \in [\![1, n+1]\!]$, soit $z_i \in G_{p_i}$. Supposons que $z_1 \cdots z_n z_{n+1} = 1_G$. On a alors

$$z_{n+1} = (z_1 \cdots z_n)^{-1}.$$

Soit d le produit des ordres de $z_1, \ldots, z_n$. Puisque G est abélien, on a

$$z_{n+1}^d = z_1^{-d} \cdots z_n^{-d} = 1_G.$$

Par ailleurs, il existe un entier $m \geqslant 1$ tel que $z_{n+1}^{p_{n+1}^m} = 1_G$. Par choix des z_i, on a $p_{n+1} \nmid d$. Ainsi, p_{n+1}^m et d sont premiers entre eux, et il existe $u, v \in \mathbb{Z}$ tels que $up_{n+1}^m + vd = 1$. On a donc

$$z_{n+1} = z_{n+1}^1 = \left(z_{n+1}^{p_{n+1}^m}\right)^u \left(z_{n+1}^d\right)^v = 1_G.$$

En reportant dans l'égalité initiale, on obtient alors $z_1 \cdots z_n = 1_G$, et par hypothèse de récurrence, on a $z_1 = \cdots = z_n = 1_G$. Ainsi, (H_{n+1}) est vraie, et cela achève la récurrence, ainsi que la démonstration. □

Un groupe fini étant un groupe de torsion, la dernière partie de la proposition 1.7 et le théorème précédent fournissent alors immédiatement le corollaire suivant.

1.9. Corollaire. *Tout groupe abélien fini est produit direct interne de ses sous-groupes de Sylow.*

2. Exposant d'un groupe

Nous allons ici introduire la notion de groupe d'exposant fini, et d'exposant d'un groupe.

2.1. Définition. Soit G un groupe quelconque. On dit que G est *d'exposant fini* s'il existe un entier $n \geqslant 1$ tel que

$$x^n = 1_G, \quad \text{pour tout } x \in G.$$

Dans ce cas, on appelle *exposant* de G le plus petit entier $n \geqslant 1$ vérifiant cette propriété. On le note $\exp(G)$.

2.2. Exemples

(1) Par définition, si G est d'exposant fini, tout élément $x \in G$ est d'ordre fini, et l'on a

$$o(x) \mid \exp(G), \quad \text{pour tout } x \in G.$$

En particulier, $\operatorname{ppcm}\big(o(x),\ x \in G\big)$ est fini, et l'on a

$$\operatorname{ppcm}\big(o(x),\ x \in G\big) \ \mid \ \exp(G).$$

On verra dans la suite que l'on a en fait égalité.

(2) Le point précédent montre en particulier qu'un groupe contenant un élément d'ordre infini n'est pas d'exposant fini.

(3) La réciproque du point (1) est fausse. Le groupe $\mathbb{Q}/\mathbb{Z}$ n'est pas d'exposant fini, bien que tous ses éléments soient d'ordre fini, puisqu'il contient des éléments de tous ordres.

(4) D'après le corollaire VI-4.15, tout groupe fini est d'exposant fini.

2.3. Lemme. *Soit G un groupe d'exposant fini. Alors,*

$$\exp(G) \ = \ \operatorname{ppcm}\big(o(x),\ x \in G\big).$$

De plus, si G est fini, on a $\exp(G) \mid |G|$.

Démonstration. Pour tout $m \in \mathbb{Z}$, on a

$$\begin{aligned} x^m = 1_G \ \text{ pour tout } x \in G \quad &\Longleftrightarrow \quad o(x) \mid m, \ \text{ pour tout } x \in G \\ &\Longleftrightarrow \quad \operatorname{ppcm}\big(o(x),\ x \in G\big) \ \mid \ m. \end{aligned}$$

Puisque G est d'exposant fini, la première équation admet au moins une solution strictement positive. Les équivalences précédentes montrent que la plus petite solution strictement positive est alors $\operatorname{ppcm}\big(o(x), x \in G\big)$, d'où l'égalité

$$\exp(G) \ = \ \operatorname{ppcm}\,\big(o(x),\ x \in G\big).$$

De plus, si G est fini, d'après le corollaire VI-4.15, $|G|$ est solution de l'équation. On en déduit immédiatement que $\exp(G) \mid |G|$. Cela achève la démonstration. □

2.4. Exemples

(1) Si G est cyclique d'ordre n, alors $\exp(G) = n$.
En effet, $\exp(G) \mid n$. Par ailleurs, G possède un élément d'ordre n, et donc $n \mid \exp(G)$, par définition du ppcm. On en déduit l'égalité désirée.

(2) On a $\exp(D_4) = 4$, puisque les ordres des rotations dans D_4 sont 1, 2 ou 4, et l'ordre des symétries dans D_4 est égal à 2.

(3) On a $\exp(\mathfrak{S}_3) = 6$, puisque les éléments de $\mathfrak{S}_3$ sont d'ordre 1, 2 ou 3.

On va maintenant s'intéresser à l'exposant d'un groupe abélien. Remarquons qu'un groupe abélien G d'exposant fini est de torsion, puisqu'alors tout élément de G est d'ordre fini divisant $\exp(G)$. En revanche, la réciproque est fausse, comme le montre déjà l'exemple du groupe $\mathbb{Q}/\mathbb{Z}$.

On continue par un résultat très classique.

2.5. Proposition. *Soit G un groupe abélien d'exposant fini. Alors, il existe un élément $x \in G$ d'ordre $\exp(G)$.*

Démonstration. Écrivons

$$\exp(G) \;=\; \operatorname{ppcm}\big(o(x),\ x \in G\big) \;=\; p_1^{m_1} \times \cdots \times p_r^{m_r},$$

où les p_i sont des premiers deux à deux distincts, et $m_i \geqslant 1$.

Soit $i \in [\![1, r]\!]$. Alors, il existe un élément $x_i \in G$ tel que $p_i^{m_i} \mid o(x_i)$. En effet, dans le cas contraire, pour tout $x \in G$, la plus grande puissance de p_i divisant $o(x)$ serait au plus $p_i^{m_i - 1}$. Mais alors, la puissance de p_i dans la décomposition de $\operatorname{ppcm}(o(x), x \in G)$ serait $< m_i$, d'où une contradiction.

Pour $i \in [\![1, r]\!]$, choisissons $x_i \in G$ un élément tel que $p_i^{m_i} \mid o(x_i)$, et écrivons $o(x_i) = p_i^{m_i} q_i$. Par le corollaire VI-4.16, on a $o(x_i^{q_i}) = p_i^{m_i}$. Comme G est abélien, une application répétée de la proposition VI-4.17 (2) montre que l'on a

$$o(x_1^{q_1} \cdots x_r^{q_r}) = p_1^{m_1} \times \cdots \times p_r^{m_r} = \exp(G),$$

d'où le résultat annoncé. □

2.6. Corollaire. *Soit G un groupe abélien fini. Alors, $\exp(G) = |G|$ si, et seulement si, G est cyclique.*

Démonstration. Si G est cyclique, l'exemple 2.4 (1) montre que l'on a l'égalité $\exp(G) = |G|$. Inversement, supposons cette égalité vérifiée. Par la proposition précédente, il existe un élément $x \in G$ tel que $o(x) = \exp(G)$. Ainsi, x est un élément de G d'ordre $|G|$, donc G est cyclique. □

2.7. Remarque. L'exemple de $\mathfrak{S}_3$ montre que ces deux derniers résultats sont faux si G n'est pas abélien.

Nous allons maintenant donner une autre application de la notion d'exposant. Avant de le faire, rappelons que pour tout corps K, on peut définir l'anneau $K[X]$ des polynômes à coefficients dans un corps K, de manière tout à fait identique à celle du cas réel ou complexe. Cet anneau de polynômes possède alors les propriétés usuelles. En particulier, on a la notion de degré d'un polynôme non nul, et tout polynôme de degré $d \geqslant 0$ a au plus d zéros dans K. Le lecteur se référera au paragraphe XV-6 pour plus de précisions.

On peut maintenant démontrer le théorème suivant.

2.8. Théorème. *Soit K un corps. Alors, tout sous-groupe fini de $K^\times$ est cyclique.*

Démonstration. Soit G un sous-groupe fini de $K^\times$. En particulier, G est abélien. Soit $e = \exp(G)$. Alors, on a $x^e = 1_K$ pour tout $x \in G$. Or, le polynôme $X^e - 1_K \in K[X]$ est de degré e, donc possède au plus e zéros dans K. Ainsi, $|G| \leqslant e$. Par ailleurs, d'après la proposition 2.5, il existe x_0 dans G d'ordre e. Comme $\langle x_0 \rangle \subset G$ et que $|G| \leqslant e$, on en déduit finalement que $|G| = e$ et que $G = \langle x_0 \rangle$. Ainsi, G est cyclique. □

Puisque $\mathbb{Z}/p\mathbb{Z}$ est un corps, pour tout nombre premier p, on obtient immédiatement le résultat suivant.

2.9. Corollaire. *Soit p un nombre premier. Alors, $(\mathbb{Z}/p\mathbb{Z})^\times$ est cyclique.*

3. Caractères linéaires d'un groupe abélien fini

3.1. Définition. Soit G un groupe abélien. On appelle *caractère linéaire* de G tout morphisme de groupes $\chi : G \longrightarrow \mathbb{C}^\times$.

L'ensemble des caractères linéaires de G est un groupe abélien pour la multiplication des fonctions. On le note $\widehat{G}$.

3.2. Remarque. Si G_1 et G_2 sont deux groupes abéliens isomorphes, alors on a un isomorphisme de groupes

$$\widehat{G}_1 \simeq \widehat{G}_2.$$

En effet, si $\varphi : G_1 \xrightarrow{\sim} G_2$ est un isomorphisme, l'application

$$\begin{aligned} \widehat{G}_2 &\longrightarrow \widehat{G}_1 \\ \chi &\longmapsto \chi \circ \varphi \end{aligned}$$

est un isomorphisme de groupes. Les détails sont laissés au lecteur en guise d'exercice.

3.3. Exemples

(1) D'après l'exercice VI-36, on a $\widehat{\mathbb{Z}} \simeq \mathbb{C}^\times$.

(2) Soit G un groupe cyclique d'ordre n. Alors, $\widehat{G}$ est aussi un groupe cyclique d'ordre n. En particulier, on a $G \simeq \widehat{G}$.
D'après la remarque précédente, il suffit de le démontrer lorsque G est le groupe $\mathbb{Z}/n\mathbb{Z}$. Par le théorème de factorisation, les morphismes de

groupes $\mathbb{Z}/n\mathbb{Z} \longrightarrow \mathbb{C}^\times$ sont en bijection avec les morphismes $\mathbb{Z} \longrightarrow \mathbb{C}^\times$ dont le noyau contient $n\mathbb{Z}$. Un tel morphisme envoie donc 1 sur un élément de $\mathbb{U}_n = \{z \in \mathbb{C} \mid z^n = 1\}$. Ainsi, les caractères linéaires du groupe $\mathbb{Z}/n\mathbb{Z}$ sont induits par passage au quotient par les morphismes de la forme

$$\begin{aligned} h_z : \mathbb{Z} &\longrightarrow \mathbb{C}^\times \\ m &\longmapsto z^m, \end{aligned}$$

avec $z \in \mathbb{U}_n$. On a donc une bijection

$$\begin{aligned} \mathbb{U}_n &\longrightarrow \widehat{G} \\ z &\longmapsto \overline{h}_z, \end{aligned}$$

et l'on vérifie que c'est en fait un isomorphisme de groupes. Comme $\mathbb{U}_n$ est cyclique d'ordre n, on a le résultat voulu.

Nous allons maintenant nous intéresser de plus près au groupe des caractères linéaires d'un groupe abélien fini. On commence par le résultat suivant.

3.4. Théorème. *Soit G un groupe abélien, et soit H un sous-groupe de G. Alors, l'application de restriction*

$$\begin{aligned} \rho_H : \widehat{G} &\longrightarrow \widehat{H} \\ \chi &\longmapsto \chi_{|H} \end{aligned}$$

est un morphisme de groupes, de noyau canoniquement isomorphe à $\widehat{G/H}$. Si de plus $[G : H]$ est fini, ρ_H est surjectif.

En particulier, si G est un groupe abélien fini, tout caractère linéaire de H se prolonge en un caractère linéaire de G.

Démonstration. Le fait que ρ_H est un morphisme de groupes est clair. Par définition, son noyau est l'ensemble des caractères linéaires de G dont le noyau contient H. Le groupe G étant abélien, H est distingué dans G, et par le théorème de factorisation, $\mathrm{Ker}(\rho_H)$ est donc en bijection avec l'ensemble des morphismes de groupes $G/H \longrightarrow \mathbb{C}^\times$, c'est-à-dire $\widehat{G/H}$. On vérifie facilement que cette bijection est un morphisme de groupes.

Supposons maintenant que $[G : H]$ soit fini. On va montrer que tout caractère linéaire de H se prolonge en un caractère linéaire de G par récurrence sur $r = [G : H]$. Plus précisément, on note (H_r) la propriété suivante :

(H_r) Pour tout sous-groupe H de G tel que $[G : H] \leqslant r$, tout caractère linéaire de H se prolonge en un caractère linéaire de G.

Si $r = 1$, alors $G = H$ et le résultat est trivial. Ainsi, (H_1) est vraie. Supposons que la propriété (H_r) soit vraie pour un entier $r \geqslant 1$, et montrons que (H_{r+1}) est vraie.

Soit H un sous-groupe de G tel que $[G : H] \leqslant r + 1$. Si $[G : H] \leqslant r$, on applique l'hypothèse de récurrence. On suppose donc que $[G : H] = r + 1$. En particulier, $[G : H] \geqslant 2$, et $H \neq G$. Fixons $g \in G \setminus H$. Soit $H' = \langle H, g \rangle$. Comme G est abélien, tout élément $h' \in H'$ s'écrit sous la forme

$$h' = hg^m, \quad h \in H,\ m \in \mathbb{Z}.$$

Soit $\chi : H \longrightarrow \mathbb{C}^\times$ un caractère linéaire de H et soit d l'ordre de $\overline{g}$ dans G/H. En particulier, $\overline{g^d} = \overline{g}^d = \overline{1}$, et donc $g^d \in H$. Soit $z \in \mathbb{C}^\times$ une racine d-ième de $\chi(g^d)$ fixée. Si $h' \in H'$ s'écrit $h' = hg^m$, on pose

$$\chi'(h') = \chi(h)z^m.$$

Cette quantité ne dépend que de h', et non pas de la décomposition choisie. En effet, si $h' = h_1 g^{m_1} = h_2 g^{m_2}$, alors on a $g^{m_2 - m_1} = h_2^{-1} h_1$, et par conséquent $\overline{g}^{m_2 - m_1} = \overline{1} \in G/H$. Ainsi, $d \mid (m_2 - m_1)$. Il existe donc $k \in \mathbb{Z}$ tel que $m_2 = m_1 + kd$. On a alors $h_1 = h_2 g^{kd}$, et ainsi

$$\begin{aligned} \chi(h_2)z^{m_2} &= \chi(h_2)(z^d)^k z^{m_1} \\ &= \chi(h_2)\chi(g^d)^k z^{m_1} \\ &= \chi(h_2 g^{kd}) z^{m_1} \\ &= \chi(h_1) z^{m_1}. \end{aligned}$$

Par conséquent, l'application $\chi' : H' \longrightarrow \mathbb{C}^\times$ est bien définie. On vérifie aisément que c'est un morphisme de groupes, qui prolonge le caractère χ. Comme $g \in G \setminus H$, H' contient strictement H, et donc $[G : H'] < [G : H]$. Ainsi, $[G : H'] \leqslant r$, et par hypothèse de récurrence, χ' se prolonge en un caractère linéaire χ'' de G. Mais alors, χ'' prolonge aussi χ, ce qui achève la récurrence.

Il s'ensuit alors que ρ_H est surjective. Le dernier point du théorème étant clair, cela achève la démonstration. □

3.5. Remarque. Nous verrons dans la dernière partie de ce chapitre que ce résultat est en fait vrai pour tout sous-groupe H.

3.6. Corollaire. *Soit G un groupe abélien fini. Alors, $|\widehat{G}| = |G|$.*

Démonstration. On procède par récurrence sur $n = |G|$. On note (H_n) la propriété suivante :

(H_n) Pour tout groupe abélien fini G tel que $|G| \leqslant n$, on a $|\widehat{G}| = |G|$.

La proposition (H_1) est clairement vraie. Supposons que (H_n) soit vraie pour un entier $n \geqslant 1$, et montrons que (H_{n+1}) est vraie. Si $|G| \leqslant n$, on applique (H_n) pour obtenir le résultat. Supposons enfin que $|G| = n+1 \geqslant 2$ et soit $g \in G \setminus \{1_G\}$. Posons $H = \langle g \rangle$. Comme H est cyclique, on a $\widehat{H} \simeq H$ par l'exemple 3.3 (2), et ainsi $|\widehat{H}| = |H|$. Comme $g \neq 1_G$, on a

$$|G/H| = \frac{|G|}{|H|} < |G|,$$

et par hypothèse de récurrence, on a $|\widehat{G/H}| = |G/H|$.

D'après le théorème 3.4, le morphisme de restriction $\rho_H : \widehat{G} \longrightarrow \widehat{H}$ est surjectif, de noyau isomorphe à $\widehat{G/H}$. D'après le premier théorème d'isomorphisme, on a $\widehat{G}/\mathrm{Ker}(\rho_H) \simeq \widehat{H}$. En particulier, on obtient

$$|\widehat{G}| = |\mathrm{Ker}(\rho_H)| \times |\widehat{H}| = |\widehat{G/H}| \times |\widehat{H}|,$$

d'où finalement

$$|\widehat{G}| = |G/H| \times |H| = |G|.$$

Cela achève la démonstration. □

3.7. Remarque. On peut montrer que pour tout groupe abélien fini, on a un isomorphisme $G \simeq \widehat{G}$ (cf. exercice 7).

3.8. Corollaire. *Soit G un groupe abélien fini. Pour tout $x \in G$, l'application*

$$\begin{aligned} ev_x : \widehat{G} &\longrightarrow \mathbb{C}^\times \\ \chi &\longmapsto \chi(x) \end{aligned}$$

est un caractère linéaire de $\widehat{G}$, et l'application

$$\begin{aligned} \theta : G &\longrightarrow \widehat{\widehat{G}} \\ x &\longmapsto ev_x \end{aligned}$$

est un isomorphisme de groupes.

Démonstration. Le seul point non immédiat est la bijectivité de θ. D'après le corollaire 3.6, on a

$$|G| = |\widehat{G}| = |\widehat{\widehat{G}}|.$$

Il suffit donc de démontrer l'injectivité de θ. Si $x \in G$ est un élément différent de 1_G, il existe au moins un caractère linéaire $\chi_0 : \langle g \rangle \longrightarrow \mathbb{C}^\times$ non trivial (car $|\widehat{\langle x \rangle}| = |\langle x \rangle| > 1$). En particulier, $\chi_0(x) \neq 1$ (sinon χ_0 serait trivial). Par le théorème 3.4, χ_0 se prolonge en un caractère linéaire χ de G. On a alors

$$\theta(x)(\chi) = \chi(x) = \chi_0(x) \neq 1,$$

et $\theta(x)$ est donc un élément non trivial de $\widehat{\widehat{G}}$. Cela démontre l'injectivité du morphisme θ, et achève la démonstration. □

Dorénavant, les groupes abéliens seront notés additivement, et ce, jusqu'à la fin de ce chapitre.

4. Groupes abéliens libres

Nous allons nous intéresser dans ce paragraphe à la notion de groupe abélien libre.

4.1. Définition. Soit G un groupe abélien. Une famille $\mathbf{e} = (e_i)_{i\in I}$ d'éléments de G est dite *libre* si pour toute famille d'entiers $(m_i)_{i\in I}$ presque tous nuls, on a

$$\sum_{i\in I} m_i{\cdot}e_i = 0 \quad \Longrightarrow \quad m_i = 0, \quad \text{pour tout } i \in I.$$

Une *base* de G est une famille à la fois libre et génératrice. Cela revient à dire que tout élément $x \in G$ s'écrit de manière **unique** sous la forme

$$x = \sum_{i\in I} m_i{\cdot}e_i, \quad m_i \in \mathbb{Z},$$

où les m_i sont presque tous nuls.

On dit que G est un *groupe abélien libre* s'il possède au moins une base.

Si X est un ensemble, un *groupe abélien libre sur* X est un couple (G, ι), où G est un groupe, et $\iota : X \hookrightarrow G$ est une application injective telle que $\iota(X)$ soit une base de G.

On dit que deux groupes abéliens libres (G, ι) et (G', ι') sur X sont *isomorphes* s'il existe un isomorphisme de groupes $\varphi : G \xrightarrow{\sim} G'$ tel que

$$\varphi\big(\iota(x)\big) = \iota'(x), \quad \text{pour tout } x \in X.$$

Attention ! Un groupe abélien libre n'est pas un groupe libre au sens de la définition IX-4.3 dès qu'une base possède au moins deux éléments, puisque une base d'un groupe abélien libre n'est pas une famille libre au sens de la définition IX-4.1.

En effet, pour tous $i, j \in I$, $i \neq j$, on a la relation $e_i + e_j - e_i - e_j = 0$.

4.2. Exemple. Soit X un ensemble, et soit $\mathbb{Z}^{(X)}$ l'ensemble des familles d'entiers $(m_x)_{x\in X}$ à support fini. Pour tout $x \in X$, on note ε_x l'élément dont toutes les coordonnées sont nulles, sauf celle qui correspond à l'indice x, qui vaut 1. Alors, $(\varepsilon_x)_{x\in X}$ est clairement une base de $\mathbb{Z}^{(X)}$, et l'application

$$\begin{aligned} \varepsilon : X &\longrightarrow \mathbb{Z}^{(X)} \\ x &\longmapsto \varepsilon_x \end{aligned}$$

est injective.

En particulier, le couple $(\mathbb{Z}^{(X)}, \varepsilon)$ est un groupe abélien libre sur X.

On peut alors obtenir une première caractérisation des groupes abéliens libres sur X.

4.3. Lemme. *Soit X un ensemble, soit G un groupe abélien, et soit*

$$\iota : X \longrightarrow G$$

une application injective. Posons $e_x = \iota(x)$, pour tout $x \in X$.

Alors, (G, ι) est un groupe abélien libre si, et seulement si, le morphisme de groupes

$$\begin{aligned} \varphi : \quad \mathbb{Z}^{(X)} &\longrightarrow G \\ (m_x)_{x\in X} &\longmapsto \sum_{x\in X} m_x \cdot e_x \end{aligned}$$

est un isomorphisme.

En particulier, tout groupe abélien libre sur X est isomorphe à $(\mathbb{Z}^{(X)}, \varepsilon)$.

Démonstration. Par définition même d'une base, la famille $\iota(X) = (e_x)_{x\in X}$ est une base de G si, et seulement si, φ est bijective, i.e. un isomorphisme. Dans ce cas, φ est un isomorphisme de groupes tel que

$$\varphi(\varepsilon_x) = e_x = \iota(x), \qquad \text{pour tout } x \in X,$$

d'où le dernier point. □

Comme dans le cas des groupes libres, les groupes abéliens libres peuvent être caractérisés comme étant les groupes vérifiant une certaine propriété universelle.

4.4. Proposition. *Soit G un groupe abélien, et soit $\iota : X \hookrightarrow G$ une application injective.*

Alors, (G, ι) est un groupe abélien libre sur X si, et seulement si, (G, ι) satisfait la propriété universelle suivante : pour tout groupe abélien G', et toute famille $\mathbf{a} = (a_x)_{x\in X}$ d'éléments de G' indexée par X, il existe un

unique morphisme de groupes $\varphi_{\mathbf{a}} : G \longrightarrow G'$ *tel que*

$$\varphi_{\mathbf{a}}\big(\iota(x)\big) = a_x, \qquad \textit{pour tout } x \in X.$$

Démonstration. Supposons que (G, ι) soit un groupe abélien libre sur X. Posons $e_x = \iota(x)$ pour tout $x \in X$. Alors, $(e_x)_{x\in X}$ est une base de G. Tout élément de G s'écrit donc sous la forme

$$x = \sum_{x\in X} m_x{\cdot}e_x, \quad m_x \in \mathbb{Z},$$

où les entiers m_x sont presque tous nuls. Soit G' un groupe abélien quelconque, et soit $\mathbf{a} = (a_x)_{x\in X}$ une famille d'éléments de G' indexée par X. Supposons qu'il existe un morphisme $\varphi_{\mathbf{a}}$ vérifiant les conditions de l'énoncé. Alors, pour toute famille d'entiers $(m_x)_{x\in X} \in \mathbb{Z}^{(X)}$, on a

$$\varphi_{\mathbf{a}}\Big(\sum_{x\in X} m_x{\cdot}e_x\Big) = \sum_{x\in X} m_x{\cdot}\varphi_{\mathbf{a}}(e_x) = \sum_{x\in X} m_x{\cdot}a_x,$$

car $\varphi_{\mathbf{a}}$ est un morphisme de groupes. Ainsi, si $\varphi_{\mathbf{a}}$ existe, il est unique et défini par la formule précédente. Il reste à vérifier que l'application

$$\begin{array}{ccc} G & \longrightarrow & G' \\ \displaystyle\sum_{x\in X} m_x{\cdot}e_x & \longmapsto & \displaystyle\sum_{x\in X} m_x{\cdot}a_x \end{array}$$

est un morphisme de groupes vérifiant les conditions demandées, ce qui est clair.

Inversement, supposons que le couple (G, ι) vérifie la propriété universelle décrite dans l'énoncé. En prenant $(G', \mathbf{a}) = \big(\mathbb{Z}^{(X)},\ (\varepsilon_x)_{x\in X}\big)$, on a donc l'existence d'un unique morphisme $\psi : G \longrightarrow \mathbb{Z}^{(X)}$ tel que

$$\psi(e_x) = \varepsilon_x, \qquad \text{pour tout } x \in X.$$

On vérifie alors sans peine que ψ et l'application φ du lemme 4.3 sont inverses l'une de l'autre. On a alors la conclusion en appliquant ce même lemme. □

Nous allons maintenant nous intéresser plus particulièrement aux groupes abéliens libres de type fini.

4.5. Théorème. *Soit G un groupe abélien libre de type fini. Alors, toutes les bases de G sont finies, et ont même cardinal.*

Démonstration. Soit $\mathbf{e} = (e_i)_{i\in I}$ une base de G et soit $(x_1, \ldots, x_r)$ une famille de générateurs de G. Chaque x_i s'écrit comme combinaison linéaire finie à coefficients entiers de certains éléments de la base. Il existe donc un

sous-ensemble $J \subset I$ fini tel que, pour tout $i \in [\![1, r]\!]$, on a

$$x_i = \sum_{j \in J} m_{ij} \cdot e_j, \quad m_{ij} \in \mathbb{Z}.$$

Comme $x_1, \ldots, x_r$ engendrent G, tout élément de G est combinaison linéaire à coefficients entiers de $x_1, \ldots, x_r$, donc aussi des éléments de la famille $(e_j)_{j \in J}$. S'il existait $i_0 \in I \setminus J$, alors e_{i_0} s'écrirait donc sous la forme

$$e_{i_0} = \sum_{j \in J} m'_j \cdot e_j, \quad m'_j \in \mathbb{Z}.$$

Comme $i_0 \notin J$, cela contredirait la liberté de la famille $\mathbf{e}$. Ainsi, $I = J$, et $\mathbf{e}$ est finie.

Il nous reste à démontrer que toutes les bases de G ont même cardinal.

Soit $(e_1, \ldots, e_r)$ une base de G. Notons

$$2 \cdot G = \{2 \cdot x \mid x \in G\}.$$

C'est un sous-groupe de G. L'application

$$\begin{array}{ccc} G & \longrightarrow & (\mathbb{Z}/2\mathbb{Z})^r \\ m_1 \cdot e_1 + \cdots + m_r \cdot e_r & \longmapsto & (\overline{m}_1, \ldots, \overline{m}_r) \end{array}$$

est surjective, de noyau $2 \cdot G$ et le théorème de factorisation nous donne un isomorphisme de groupes

$$G/2 \cdot G \simeq (\mathbb{Z}/2\mathbb{Z})^r.$$

Si $(e'_1, \ldots, e'_s)$ est une autre base de G, on a de même $G/2 \cdot G \simeq (\mathbb{Z}/2\mathbb{Z})^s$, et par suite, on obtient un isomorphisme

$$(\mathbb{Z}/2\mathbb{Z})^r \simeq (\mathbb{Z}/2\mathbb{Z})^s.$$

Mais, deux groupes isomorphes ayant même ordre, on en déduit que $r = s$, d'où le résultat. □

4.6. Définition. Soit G un groupe abélien libre de type fini. Le cardinal de n'importe quelle base de G est appelé le *rang* de G, et est noté $\mathrm{rg}(G)$.

4.7. Remarque. On peut démontrer plus généralement que toutes les bases d'un groupe abélien libre sont équipotentes. Le lecteur se référera par exemple à [9, Théorème 2.11] pour une démonstration (en appliquant le résultat à $A = \mathbb{Z}$).

Les résultats démontrés jusqu'à présent permettent d'obtenir immédiatement le résultat suivant.

4.8. Proposition. *Tout groupe abélien libre de rang r est isomorphe à $\mathbb{Z}^r$. En particulier, deux groupes abéliens libres de type fini sont isomorphes si, et seulement si, il ont même rang.*

Nous allons maintenant nous intéresser au sous-groupe de torsion d'un groupe abélien libre. On a le théorème important suivant.

4.9. Théorème. *Tout groupe abélien libre est sans torsion. De plus, un groupe abélien G de type fini est libre si, et seulement s'il est sans torsion. Dans ce cas, si G est engendré par n éléments, on a $rg(G) \leqslant n$.*

Démonstration. Soit G un groupe abélien. Supposons que G soit libre, et soit $\mathbf{e} = (e_x)_{x\in X}$ une base de G. Soit $x = \sum_{x\in X} m_x \cdot e_x \in G$. Supposons qu'il existe un entier $m \geqslant 1$ tel que $m\cdot x = 0$. Alors, on a

$$\sum_{x\in X} (mm_x)\cdot e_x = 0,$$

et comme $\mathbf{e}$ est libre, on obtient

$$mm_x = 0, \quad \text{pour tout } x \in X,$$

d'où $m_x = 0$ pour tout $x \in X$, puisque $m \geqslant 1$. Ainsi, $x = 0$ et G est sans torsion.

Supposons maintenant que G soit un groupe abélien de type fini. Il reste à démontrer que si $\mathrm{Tor}(G) = \{0\}$, alors G est libre. Nous allons le faire par récurrence. Soit $(x_1, \ldots, x_n)$ une famille génératrice de G.

Si $n = 1$, alors G est monogène. Si $x_1 = 0$, G est nul, donc libre de base vide, et de rang nul. Si x_1 est non nul, comme x_1 n'est pas d'ordre fini par hypothèse, G est un groupe monogène infini, donc isomorphe à $\mathbb{Z}$ par le théorème VI-6.3, et le résultat est clair. Supposons alors que tout groupe abélien sans torsion engendré par n éléments soit libre de rang $\leqslant n$, et supposons que G soit engendré par $n+1$ éléments $x_1, \ldots, x_{n+1}$. Si un des x_i est nul, on peut appliquer l'hypothèse de récurrence. Supposons maintenant que tous les x_i soient non nuls. Si la famille $(x_1, \ldots, x_{n+1})$ est libre, c'est fini. Sinon, il existe des entiers $(m_1, \ldots, m_{n+1})$ non tous nuls tels que

$$m_1\cdot x_1 + \cdots + m_{n+1}\cdot x_{n+1} = 0.$$

Il y a au moins deux coefficients de la relation qui sont non nuls. Sinon, on aurait une relation de coefficient $m_i x_i = 0$, avec $m_i \neq 0$, d'où la contradiction $x_i = 0$ puisque G est sans torsion. Quitte à renuméroter les générateurs,

on peut supposer que

$$|m_1| \geqslant |m_2| \geqslant 1.$$

Écrivons $m_1 = m_2 q + r$, $q, r \in \mathbb{Z}$, $r \in [\![0, |m_2| - 1]\!]$. En particulier, on a

$$r < |m_2| \leqslant |m_1|.$$

On obtient alors

$$r{\cdot}x_1 + m_2{\cdot}(x_2 + qx_1) + m_3 x_3 + \cdots + m_{n+1}{\cdot}x_{n+1} = 0.$$

Remarquons maintenant que cette nouvelle relation a encore au moins un coefficient non nul, à savoir m_2. De plus, on a

$$|r| + |m_2| + \cdots + |m_{n+1}| < |m_1| + |m_2| + \cdots + |m_{n+1}|,$$

et la famille $(x_1, x_2 + q{\cdot}x_1, x_3, \ldots, x_{n+1})$ est encore génératrice. À partir d'une famille génératrice et d'une relation de dépendance ayant au moins deux coefficients non nuls, on peut donc construire une nouvelle famille génératrice à $n+1$ éléments vérifiant une relation de dépendance avec au moins un coefficient non nul, et dont la somme des valeurs absolues des coefficients est strictement inférieure.

Au bout d'un nombre fini d'étapes, on aura alors construit une famille génératrice $(x'_1, \ldots, x'_{n+1})$ de G vérifiant une relation de dépendance de la forme $m_i{\cdot}x'_i = 0$ avec $m_i \neq 0$. Comme G est sans torsion, $x'_i = 0$. Mais alors, G peut être engendré par n éléments, et l'on applique l'hypothèse de récurrence. Cela achève la démonstration. □

4.10. Remarque. Ce théorème est faux si G n'est pas de type fini, puisque $\mathbb{Q}$ est abélien, sans torsion, mais n'est pas libre.

Remarquons que deux éléments non nuls de $\mathbb{Q}$ sont toujours liés. En effet, si $\frac{a}{b}, \frac{c}{d} \in \mathbb{Q}$, on a

$$bc\frac{a}{b} - ad\frac{c}{d} = 0.$$

En particulier, si $\mathbb{Q}$ possède une base, elle a au plus un élément. Mais alors, il existerait $u, v \in \mathbb{Z}$ premiers entre eux tels que $\mathbb{Q} = \frac{u}{v}\mathbb{Z}$. On obtient une contradiction en considérant un nombre premier p ne divisant pas v, et en remarquant que $\frac{1}{p} \notin \frac{u}{v}\mathbb{Z}$.

On s'intéresse maintenant aux sous-groupes d'un groupe libre.

4.11. Corollaire. *Soit G un groupe abélien libre de rang r. Alors, tout sous-groupe de G est abélien libre de rang $\leqslant r$.*

Démonstration. Soit G un groupe abélien libre de rang r, et soit H un sous-groupe de G. Alors, G étant sans torsion d'après le théorème 4.9, le sous-groupe H est abélien et sans torsion. Si l'on montre que H peut

être engendré par au plus r éléments, on aura la conclusion souhaitée en appliquant ce même théorème.

On procède par récurrence sur r. Soit (H_r) la propriété suivante :

(H_r) Tout sous-groupe d'un groupe abélien libre de rang r peut être engendré par au plus r éléments.

Si $r = 0$, le résultat est clair, puisque qu'un groupe abélien libre de rang nul est le groupe trivial. Ainsi, (H_0) est vraie. Supposons avoir montré (H_r) pour un certain entier $r \geqslant 0$ et montrons que (H_{r+1}) est vraie.

Soit G un groupe abélien libre de rang $r+1$, soit $\mathbf{e} = (e_1, \ldots, e_{r+1})$ une base de G et soit H un sous-groupe de G.

Pour tout $h \in H$, on écrit

$$h = m_{1h} \cdot e_1 + \cdots + m_{(r+1)h} \cdot e_{r+1}, \quad m_{ih} \in \mathbb{Z}.$$

Soit $A = \langle m_{1h}, h \in H \rangle \subset \mathbb{Z}$. Par la proposition VI-2.4, il existe $d_1 \in \mathbb{Z}$ tel que $A = d_1\mathbb{Z}$. On a donc

$$m_{1h} = d_1 m'_{1h}, \quad m'_{1h} \in \mathbb{Z}, \quad \text{pour tout } h \in H.$$

Comme $d_1 \in d_1\mathbb{Z} = A$, il existe des entiers $n_h \in \mathbb{Z}$ presque tous nuls tels que

$$d_1 = \sum_{h \in H} n_h m_{1h}.$$

Posons $y_1 = \sum_{h \in H} n_h h$. On a

$$y_1 = \sum_{h \in H} \sum_{i=1}^{r+1} n_h m_{ih} \cdot e_i = \sum_{i=1}^{r+1} \Big(\sum_{h \in H} n_h m_{i_h} \Big) \cdot e_i = d_1 \cdot e_1 + z_1,$$

où $z_1 \in G$ est combinaison linéaire de $e_2, \ldots, e_{r+1}$. Autrement dit, y_1 est un élément de H dont la coordonnée en e_1 est égale à d_1. Mais alors, on a

$$H = \langle h \in H \rangle = \langle y_1,\ h \in H \rangle = \langle y_1,\ h - m'_{1h} \cdot y_1,\ h \in H \rangle.$$

Le sous-groupe $H' = \langle h - m'_{1h} \cdot y_1,\ h \in H \rangle$ est alors un sous-groupe du groupe abélien libre $G' = \langle e_2, \ldots, e_{r+1} \rangle$, lequel est de rang r. Par hypothèse de récurrence, H' peut être engendré par au plus r éléments, et H peut donc être engendré par au plus $r+1$ éléments, ce qui achève la démonstration. □

5. Structure des groupes abéliens de type fini

Nous allons ici donner un théorème de structure pour les groupes abéliens de type fini. Comme nous gardons une notation additive, le coproduit direct

interne de groupes sera noté $\oplus$, et appelé *somme directe interne.* Le lemme suivant permet un premier dévissage.

5.1. Lemme. *Soit G un groupe abélien de type fini. Alors,* $\mathrm{Tor}(G)$ *est fini, et il existe un sous-groupe abélien libre de type fini L tel que*

$$G = L \oplus \mathrm{Tor}(G).$$

Démonstration. Remarquons que $G/\mathrm{Tor}(G)$ est abélien, de type fini, et sans torsion par le lemme 1.4. Ainsi, $G/\mathrm{Tor}(G)$ est un groupe abélien libre de type fini d'après le théorème 4.9.

Soit $(\overline{e}_1, \ldots, \overline{e}_n)$ une base de $G/\mathrm{Tor}(G)$. Montrons que $(e_1, \ldots, e_n)$ est libre. Soient $m_1, \ldots, m_n \in \mathbb{Z}$ des entiers tels que

$$m_1 \cdot e_1 + \cdots + m_n \cdot e_n = 0.$$

Alors, on a

$$m_1 \cdot \overline{e}_1 + \cdots + m_1 \cdot \overline{e}_n \;=\; \overline{m_1 \cdot e_1 + \cdots + m_n \cdot e_n} \;=\; \overline{0},$$

d'où $m_1 = \cdots = m_n = 0$, car $(\overline{e}_1, \ldots, \overline{e}_n)$ est libre. Le sous-groupe

$$L = \langle e_1, \ldots, e_n \rangle$$

est donc un sous-groupe de G abélien libre de type fini.

Montrons que $G = L \oplus \mathrm{Tor}(G)$. Soit $x \in G$. Puisque $(\overline{e}_1, \ldots, \overline{e}_n)$ est une base de $G/\mathrm{Tor}(G)$, il existe $m_1, \ldots, m_n \in \mathbb{Z}$ tel que

$$\overline{x} \;=\; m_1 \cdot \overline{e}_1 + \cdots + m_1 \cdot \overline{e}_n \;=\; \overline{m_1 \cdot e_1 + \cdots + m_n \cdot e_n}.$$

Mais alors $x = (m_1 \cdot e_1 + \cdots + m_n \cdot e_n) + t$, pour un certain $t \in \mathrm{Tor}(G)$, d'où l'existence d'une décomposition. Pour montrer l'unicité, il suffit de vérifier que $L \cap \mathrm{Tor}(G) = \{0\}$. Soit $x \in L \cap \mathrm{Tor}(G)$. Comme $x \in L$, on peut écrire

$$x = m_1 \cdot e_1 + \cdots + m_n \cdot e_n, \quad m_i \in \mathbb{Z}.$$

Puisque $x \in \mathrm{Tor}(L)$, il existe $m \geqslant 1$ tel que

$$m \cdot x = m m_1 \cdot e_1 + \cdots + m m_n \cdot e_n = 0.$$

Comme la famille $(e_1, \ldots, e_n)$ est libre, on a $m m_i = 0$, pour tout $i \in [\![1, n]\!]$, et donc $m_i = 0$, puisque $m \geqslant 1$.

On a donc bien montré que $G = L \oplus \mathrm{Tor}(G)$.

Il reste à démontrer que $\mathrm{Tor}(G)$ est fini. Puisque $G = L \oplus \mathrm{Tor}(G)$, la projection sur $\mathrm{Tor}(G)$ induit un isomorphisme de groupes

$$G/L \simeq \mathrm{Tor}(G).$$

En particulier, le groupe abélien $\mathrm{Tor}(G)$ est de type fini. Soient $x_1, \ldots, x_r$ des générateurs de $\mathrm{Tor}(G)$. Tout élément de $x \in \mathrm{Tor}(G)$ s'écrit donc

$$x = m_1 \cdot x_1 + \cdots + m_r \cdot x_r, \quad m_i \in \mathbb{Z}.$$

Si $n_i = o(x_i)$, on a

$$\langle x_i \rangle = \{0, x_i, \ldots, (n_i - 1) \cdot x_i\}.$$

Comme $m_i \cdot x_i \in \langle x_i \rangle$, on voit que x s'écrit

$$x = m_1' \cdot x_1 + \cdots + m_r' \cdot x_r, \quad m_i' \in [\![0, n_i - 1]\!].$$

Par conséquent, $|\mathrm{Tor}(G)| \leqslant n_1 \cdots n_r$. Cela achève la démonstration. □

5.2. Remarque. Ce résultat est faux si G n'est pas de type fini. Le lecteur trouvera un contre-exemple dans l'exercice 13.

Ce lemme permet de réduire l'étude des groupes abéliens de type fini à l'étude de deux cas particuliers, à savoir les groupes abéliens libres de type fini, et les groupes abéliens finis. Le premier cas a été déjà étudié en détail dans le paragraphe précédent. Nous allons maintenant nous intéresser au second.

5.3. Théorème. *Soit G un groupe abélien fini.*

Alors, il existe des entiers $d_1, \ldots, d_s \geqslant 2$ vérifiant $d_1 \mid d_2 \mid \cdots \mid d_s$ tels que

$$G \simeq \mathbb{Z}/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_s\mathbb{Z}.$$

De plus, la suite d'entiers $(d_1, \ldots, d_s)$ est unique, et ne dépend que de la classe d'isomorphisme de G.

Démonstration. Soit G un groupe abélien fini. On va procéder par récurrence sur $n = |G|$. Soit (H_n) la propriété suivante :

(H_n) La partie existence du théorème est vraie pour tout groupe abélien fini G tel que $|G| \leqslant n$.

Si $n = 1$, alors G est le groupe trivial, et le résultat est vrai. Supposons avoir montré (H_n) pour un entier $n \geqslant 1$, et montrons que (H_{n+1}) est vraie.

Soit G un groupe abélien d'ordre $|G| \leqslant n + 1$. On peut toujours supposer que $|G| = n + 1 \geqslant 2$ (dans le cas contraire, on applique l'hypothèse de récurrence). Soit $e = \exp(G)$. D'après la proposition 2.5, il existe un élément $x_0 \in G$ d'ordre e. Puisque $|G| = n + 1 \geqslant 2$, G n'est pas trivial, et l'on a $e \geqslant 2$. Comme $\langle x_0 \rangle$ et $\mathbb{U}_e = \{z \in \mathbb{C} \mid z^e = 1\}$ sont deux groupes cycliques d'ordre e, il existe un isomorphisme

$$\chi_0 : \langle x_0 \rangle \xrightarrow{\sim} \mathbb{U}_e,$$

d'après le théorème VI-6.3. Ce morphisme χ_0 étant un caractère linéaire

de $\langle x_0 \rangle$, il se prolonge en un caractère linéaire $\chi : G \longrightarrow \mathbb{C}^\times$, d'après le théorème 3.4. Remarquons que $\mathbb{U}_e \subset \mathrm{Im}(\chi)$, puisque χ prolonge χ_0. Par ailleurs, pour tout $x \in G$, on a $e \cdot x = 0$, et donc

$$\chi(x)^e = \chi(e \cdot x) = 1.$$

Ainsi, $\mathrm{Im}(\chi) \subset \mathbb{U}_e$ et par conséquent $\mathrm{Im}(\chi) = \mathbb{U}_e$. On a donc un morphisme surjectif

$$\chi : G \longrightarrow \mathbb{U}_e.$$

La composée de l'isomorphisme $\chi_0^{-1} : \mathbb{U}_e \xrightarrow{\sim} \langle x_0 \rangle$ et de l'inclusion $\langle x_0 \rangle \subset G$ fournit un morphisme $s : \mathbb{U}_e \longrightarrow G$ tel que $\chi \circ s = \mathrm{Id}_{\mathbb{U}_e}$. Autrement dit, si l'on note $H = \mathrm{Ker}(\chi)$, on a une suite exacte scindée

$$\{0\} \longrightarrow H \hookrightarrow G \underset{s}{\overset{\chi}{\rightleftarrows}} \mathbb{U}_e \longrightarrow \{0\}.$$

Par le corollaire X-3.16, on a

$$G \simeq H \times \mathbb{U}_e \simeq H \times \mathbb{Z}/e\mathbb{Z},$$

le second isomorphisme provenant du fait que $\mathbb{U}_e \simeq \mathbb{Z}/e\mathbb{Z}$. Le premier théorème d'isomorphisme appliqué à χ montre que l'on a

$$G/H \simeq \mathbb{U}_e.$$

En particulier,

$$|H| = \frac{|G|}{e} < |G|.$$

Par hypothèse de récurrence, il existe $s \geqslant 0$ et des entiers $d_1, \ldots, d_s \geqslant 2$ vérifiant $d_1 \mid \cdots \mid d_s$ tels que

$$H \simeq \mathbb{Z}/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_s\mathbb{Z}.$$

Mais alors, on a

$$G \simeq \mathbb{Z}/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_s\mathbb{Z} \times \mathbb{Z}/e\mathbb{Z}.$$

Il reste à vérifier que $d_s \mid e$. Le groupe $\mathbb{Z}/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_s\mathbb{Z} \times \mathbb{Z}/e\mathbb{Z}$ possède un élément d'ordre d_s, à savoir $(\overline{0}, \ldots, \overline{0}, \overline{1}, \overline{0})$. Mais alors, G possède un élément d'ordre d_s. En particulier, d_s divise le ppcm des ordres des éléments de G, qui est égal à e par le lemme 2.3.

Cela achève l'existence de la décomposition voulue, par récurrence.

Nous allons maintenant proposer deux démonstrations de la partie unicité du théorème.

La première est une démonstration qui n'utilise que des outils élémentaires. La seconde, quant à elle, est relativement plus courte, et s'appuie sur le théorème IX-2.7.

Première démonstration. Notons tout d'abord que pour tous $n, m \geqslant 1$, le groupe $n \cdot \mathbb{Z}/m\mathbb{Z} = \langle \overline{n} \rangle$ est cyclique d'ordre $\dfrac{m}{\mathrm{pgcd}(n,m)}$.

En effet, puisque $\overline{n} = n\cdot\overline{1}$, par le corollaire VI-4.16, on a donc

$$|n\cdot\mathbb{Z}/m\mathbb{Z}| = \frac{m}{\operatorname{pgcd}(n,m)}.$$

Supposons que l'on ait un isomorphisme

$$G \simeq \mathbb{Z}/d_1\mathbb{Z}\times\cdots\times\mathbb{Z}/d_s\mathbb{Z} \simeq \mathbb{Z}/d'_1\mathbb{Z}\times\cdots\times\mathbb{Z}/d'_t\mathbb{Z},$$

avec des notations évidentes. On peut supposer que $s \geqslant t$. Si $s = 0$, on a alors $t = 0$, et il n'y a rien à faire.

Supposons que $s \geqslant 1$. On a donc $d_1G \simeq d_1\cdot\mathbb{Z}/d_1\mathbb{Z}\times\cdots\times d_1\cdot\mathbb{Z}/d_s\mathbb{Z}$, et ainsi

$$|d_1G| = \frac{d_1\times\cdots\times d_s}{d_1^s} = \frac{|G|}{d_1^s}.$$

Mais, on a aussi $d_1G \simeq d_1\cdot\mathbb{Z}/d'_1\mathbb{Z}\times\cdots\times d_1\cdot\mathbb{Z}/d'_t\mathbb{Z}$, et ainsi

$$|d_1G| = \frac{|G|}{\operatorname{pgcd}(d_1,d'_1)\times\cdots\times\operatorname{pgcd}(d_1,d'_t)}.$$

On a donc

$$d_1^s = \operatorname{pgcd}(d_1,d'_1)\times\cdots\times\operatorname{pgcd}(d_1,d'_t) \leqslant d_1^t.$$

Comme $s \geqslant 1$, on a $d_1 \geqslant 2$. Si l'on avait $s > t$, l'inégalité ne pourrait alors avoir lieu, et donc $s = t$. Si l'on avait $\operatorname{pgcd}(d_1,d'_1) < d_1$, le membre de droite serait $< d_1^s$. Ainsi, on a $\operatorname{pgcd}(d_1,d'_1) = d_1$, soit encore $d_1 \mid d'_1$. Puisque $s = t$, on peut échanger les rôles des d_i et des d'_i. On obtient cette fois $d'_1 \mid d_1$, d'où $d_1 = d'_1$. En particulier, d'_1 divise d_i pour tout $i \in [\![1,s]\!]$.

Puisque $d_1 \mid d_2$, on a $d_2\mathbb{Z}/d_1\mathbb{Z} = 0$. On a ainsi

$$d_2G \simeq d_2\mathbb{Z}/d_2\mathbb{Z}\times\cdots\times d_2\mathbb{Z}/d_s\mathbb{Z}.$$

Mais, puisque $d'_1 \mid d_2$, on a aussi $d_2\mathbb{Z}/d'_1\mathbb{Z} = 0$, et ainsi

$$d_2G \simeq d_2\mathbb{Z}/d'_2\mathbb{Z}\times\cdots\times d_2\mathbb{Z}/d'_s\mathbb{Z}.$$

En appliquant le même raisonnement que précédemment, on obtient l'égalité $d_2 = d'_2$, et de proche en proche, on a

$$d_i = d'_i, \quad \text{pour tout } i \in [\![1,s]\!].$$

Seconde démonstration. Nous allons démontrer l'unicité par récurrence sur l'ordre de G. Plus précisément, pour tout $n \geqslant 1$, soit (H_n) la propriété

(H_n) Pour tout groupe abélien fini G tel que $|G| \leqslant n$, si l'on a

$$G \simeq \mathbb{Z}/d_1\mathbb{Z}\times\cdots\times\mathbb{Z}/d_s\mathbb{Z} \simeq \mathbb{Z}/d'_1\mathbb{Z}\times\cdots\times\mathbb{Z}/d_t\mathbb{Z},$$

avec $s, t \geqslant 0$, $d_i, d'_j \geqslant 2$, $d_1 \mid \cdots \mid d_s$ et $d'_1 \mid \cdots \mid d'_t$, alors $s = t$ et $d_i = d'_i$, pour tout $i \in [\![1,r]\!]$.

Si G est trivial, on a nécessairement $s = t = 0$ et il n'y a rien à faire. Par conséquent, (H_1) est vraie. Supposons que (H_n) soit vraie pour un certain entier $n \geqslant 1$, et montrons que (H_{n+1}) est vraie. Soit G un groupe abélien fini d'ordre $|G| \leqslant n+1$. On peut toujours supposer que G est d'ordre $n+1$ (sinon, on applique l'hypothèse de récurrence et l'on a fini). En particulier, on a $|G| \geqslant 2$, et G est non trivial. Gardons les notations précédentes. Un calcul élémentaire montre que l'on a

$$d_s \cdot (\mathbb{Z}/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_s\mathbb{Z}) = 0.$$

Ainsi, tout élément de $\mathbb{Z}/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_s\mathbb{Z}$ est d'ordre divisant d_s. De plus, l'élément $(\overline{0}, \ldots, \overline{0}, \overline{1})$ est d'ordre d_s. On en déduit que

$$\exp(\mathbb{Z}/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_s\mathbb{Z}) = d_s.$$

De même, $\exp(\mathbb{Z}/d'_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d'_t\mathbb{Z}) = d'_t$. Comme deux groupes isomorphes ont des exposants égaux, on a finalement

$$\exp(G) = d_s = d'_t.$$

Posons $e = \exp(G)$. Puisque G est non trivial, on a $e \geqslant 2$. On a donc

$$G \simeq G_1 \times \mathbb{Z}/e\mathbb{Z} \simeq G_2 \times \mathbb{Z}/e\mathbb{Z},$$

avec $G_1 = \mathbb{Z}/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_{s-1}\mathbb{Z}$ et $G_2 = \mathbb{Z}/d'_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d'_{t-1}\mathbb{Z}$. D'après le théorème IX-2.7, on obtient $G_1 \simeq G_2$. Puisque $e \geqslant 2$, on a $|G_1| < |G|$, et ainsi $|G_1| \leqslant n$. L'hypothèse de récurrence appliquée au groupe G_1 montre que l'on a $s - 1 = t - 1$, soit $s = t$, et $d_i = d'_i$, pour tout $i \in [\![1, s-1]\!]$. Comme on a par ailleurs $d_s = d'_s = e$, la propriété (H_{n+1}) est vraie, et cela achève la récurrence.

Le fait que $(d_1, \ldots, d_s)$ ne dépende que de la classe d'isomorphisme de G découle de la partie unicité. Cela achève la démonstration du théorème. □

5.4. Définition. Soit G un groupe abélien fini. Les entiers $d_1, \ldots, d_s$ fournis par le théorème précédent sont appelés les *invariants de similitude* de G.

Au vu du théorème de structure, le résultat suivant est clair.

5.5. Corollaire. *Deux groupes abéliens finis sont isomorphes si, et seulement si, ils ont mêmes invariants de similitude.*

5.6. Corollaire. *Soit G un groupe abélien fini. Alors, on a un isomorphisme de groupes*

$$G \simeq \prod_p \prod_{k \geqslant 1} \mathbb{Z}/p^{n_{p,k}}\mathbb{Z},$$

où p parcourt l'ensemble des nombres premiers, les entiers $n_{p,k}$ sont presque tous nuls, et $(n_{p,k})_{k \geqslant 1}$ est une suite d'entiers décroissante pour tout p.

De plus, les entiers $n_{p,k}$ sont uniques, et ne dépendent que de la classe d'isomorphisme de G.

Démonstration. Commençons par remarquer que pour tous entiers $n, m \geqslant 1$ premiers entre eux, on a un isomorphisme [1]

$$\mathbb{Z}/nm\mathbb{Z} \simeq \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}.$$

Une manière expéditive de le montrer est de remarquer que $(\bar{1}, \tilde{1})$ est d'ordre nm dans $\mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/m\mathbb{Z}$, et donc que ce dernier groupe est cyclique d'ordre nm. En particulier, si $p_1, \ldots, p_r$ sont des nombres premiers distincts, alors pour tous $m_1, \ldots, m_r \geqslant 0$, on a

$$\mathbb{Z}/(p_1^{m_1} \times \cdots \times p_r^{m_r})\mathbb{Z} \simeq \mathbb{Z}/p_1^{m_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_r^{m_r}\mathbb{Z}.$$

Si maintenant G est un groupe abélien fini, en décomposant les invariants de similitude de G en produits de puissances de nombres premiers, en utilisant l'isomorphisme précédent et en réarrangeant les termes, on obtient l'existence [2] de la décomposition voulue.

Montrons maintenant l'unicité des entiers $n_{p,k}$. Puisque ces entiers sont presque tous nuls, soit s le plus grand entier tel qu'il existe au moins un entier $n_{p,s}$ non nul. Alors, on a

$$G \simeq \prod_p \prod_{k=1}^{s} \mathbb{Z}/p^{n_{p,k}}\mathbb{Z} \simeq \prod_{k=1}^{s} \prod_p \mathbb{Z}/p^{n_{p,k}}\mathbb{Z},$$

le dernier isomorphisme provenant du fait qu'il n'y a qu'un nombre fini de termes.

Pour tout $k \in [\![1, s]\!]$, on pose $d_{s+1-k} = \prod_p p^{n_{p,k}}$. Comme, à p fixé, la suite d'entiers $(n_{p,k})_{k \geqslant 1}$ est décroissante, on a $d_1 \mid \cdots \mid d_s$. Remarquons que l'on a $d_1 \geqslant 2$ dès que $s \geqslant 1$, et donc que $d_k \geqslant 2$, pour tout $k \in [\![1, s]\!]$, dans ce cas.

En utilisant les isomorphismes décrits plus haut, on obtient

$$G \simeq \mathbb{Z}/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_s\mathbb{Z}.$$

Les entiers $d_1, \ldots, d_s$ étant uniques par le théorème 5.3, les entiers $n_{p,k}$ aussi, par unicité de la décomposition en facteurs premiers. Cela achève la démonstration. □

5.7. Remarque. Il est bon de savoir passer d'une décomposition à l'autre.

1. On reviendra dans un chapitre ultérieur sur cet isomorphisme, que l'on peut en fait rendre explicite.

2. On aurait aussi pu utiliser le corollaire 1.9 et appliquer le théorème 5.3 à chacun des p-Sylow.

(1) Si $G \simeq \mathbb{Z}/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_s\mathbb{Z}$, on décompose chaque d_i en produit de puissance de nombres premiers, et l'on utilise l'isomorphisme

$$\mathbb{Z}/(p_1^{m_1} \times \cdots \times p_r^{m_r})\mathbb{Z} \simeq \mathbb{Z}/p_1^{m_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_r^{m_r}\mathbb{Z}.$$

Après réarrangement des termes, on obtient une décomposition de G en produit direct de groupes de la forme $\mathbb{Z}/p^n\mathbb{Z}$, p premier, $n \geqslant 1$.

(2) Si $G \simeq \prod_p \prod_{k \geqslant 1} \mathbb{Z}/p^{n_{p,k}}\mathbb{Z}$, on choisit le plus grand entier $s \geqslant 0$ tel qu'il existe au moins un entier $n_{p,s}$ non nul. Pour tout $k \in [\![1, r]\!]$, on pose alors $d_{s+1-k} = \prod_p p^{n_{p,k}}$, et l'on obtient

$$G \simeq \mathbb{Z}/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_s\mathbb{Z}.$$

5.8. Exemples

(1) Soit $G = \mathbb{Z}/24\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$. Alors,

$$\mathbb{Z}/24\mathbb{Z} \simeq \mathbb{Z}/8\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z},$$

et donc

$$G \simeq \mathbb{Z}/2\mathbb{Z} \times (\mathbb{Z}/3\mathbb{Z})^3 \times \mathbb{Z}/8\mathbb{Z}.$$

(2) Soit $G = \mathbb{Z}/4\mathbb{Z} \times (\mathbb{Z}/2\mathbb{Z})^2 \times \mathbb{Z}/9\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z}$.

Ici, on a trois facteurs contribuant pour les puissances de 2, deux facteurs pour les puissances de 3, et un facteur pour les puissances de 5. Les autres nombres premiers n'interviennent pas. On a donc $s = 3$. Pour plus de visibilité, on écrit les puissances de chaque nombre premier dans un tableau de taille $s \times s$, en ordre décroissant. Les cases vides correspondent par convention à des puissances triviales (i.e. à des 1) :

$$\begin{array}{ccc} 4 & 2 & 2 \\ 9 & 3 & \\ 5 & & \end{array}.$$

L'entier d_k est obtenu en faisant le produit des éléments de la k-ième colonne en partant de la droite.

On a donc

$$d_1 = 2, \quad d_2 = 6, \quad d_3 = 180,$$

et ainsi

$$G \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/6\mathbb{Z} \times \mathbb{Z}/180\mathbb{Z}.$$

On peut maintenant énoncer et démontrer le théorème suivant.

5.9. Théorème (Structure des groupes abéliens de type fini) *Soit G un groupe abélien de type fini. Alors, il existe des entiers r, $s \geqslant 0$*

et des entiers $d_1, \dots, d_s \geqslant 2$ *vérifiant* $d_1 \mid \cdots \mid d_s$ *tels que*

$$G \simeq \mathbb{Z}^r \times \mathbb{Z}/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_s\mathbb{Z}.$$

De plus, l'entier r *et la suite d'entiers* $(d_1, \dots, d_s)$ *sont uniques.*

Démonstration. Par le lemme 5.1, on a

$$G = L \oplus \mathrm{Tor}(G),$$

où L est abélien libre de type fini. On a donc

$$G \simeq L \times \mathrm{Tor}(G).$$

On applique alors la proposition 4.8 et le théorème 5.3 pour avoir l'existence d'une décomposition.

Pour la partie unicité, supposons d'abord que l'on ait des isomorphismes

$$G \simeq \mathbb{Z}^{r_1} \times G_1 \simeq \mathbb{Z}^{r_2} \times G_2,$$

où G_1, G_2 sont finis. On remarque que l'isomorphisme

$$G \simeq \mathbb{Z}^{r_i} \times G_i$$

se restreint en un isomorphisme

$$\mathrm{Tor}(G) \simeq \mathrm{Tor}(\mathbb{Z}^{r_i} \times G_i) = \{0\} \times G_i,$$

et l'on a ainsi $G/\mathrm{Tor}(G) \simeq \mathbb{Z}^{r_i}$. Par conséquent, $r_1 = r_2 = \mathrm{rg}\big(G/\mathrm{Tor}(G)\big)$. Remarquons que l'on a aussi $G_1 \simeq \mathrm{Tor}(G) \simeq G_2$. Mais alors, l'unicité de la suite $(d_1, \dots, d_s)$ découle du théorème 5.3. □

6. Groupes abéliens divisibles

Nous allons ici nous intéresser à une autre famille de groupes abéliens, à savoir les groupes abéliens divisibles. Encore une fois, nous noterons les groupes abéliens de manière additive (sauf éventuellement dans certains exemples).

Notation. Pour tout groupe abélien G, et tout entier $m \in \mathbb{Z}$, on pose

$$m{\cdot}G = \{m{\cdot}x \mid x \in G\}.$$

Remarquons que $m{\cdot}G$ est un sous-groupe de G.

Lorsque G est noté multiplicativement, on le note $G^m = \{x^m \mid x \in G\}$.

6.1. Définition. Soit $n \geqslant 1$ un entier. On dit qu'un groupe abélien G est *divisible* si $G = n{\cdot}G$, pour tout $n \geqslant 1$.

6.2. Exemples

(1) Le groupe $\mathbb{Q}$ est divisible. Plus généralement, tout $\mathbb{Q}$-espace vectoriel est divisible.

(2) Le groupe $\mathbb{C}^\times$ est divisible.

(3) Soit G est un groupe abélien d'exposant fini. Alors, G est divisible si, et seulement si, il est trivial. En particulier, le seul groupe abélien fini divisible est le groupe trivial.

En effet, si G est divisible, on a $G = \exp(G)\cdot G = \{0\}$.

(4) Le groupe $\mathbb{R}^\times$ n'est pas divisible, puisque $-1 \notin \mathbb{R}^{\times 2}$.

(5) Le groupe $\mathbb{R}_+^\times$ est divisible.

On résume ci-après les propriétés de base des groupes abéliens divisibles.

6.3. Lemme. *Soient $(G_i)_{i\in I}$ une famille de groupes abéliens, soit G un groupe abélien, soit $(H_i)_{i\in I}$ une famille de sous-groupes de G.*

(1) *Si G est divisible, alors* $\mathrm{Tor}(G)$ *est divisible.*

(2) *Soit $f : G \longrightarrow G'$ un morphisme de groupes. Si G est divisible, alors le groupe* $\mathrm{Im}(f)$ *est divisible. En particulier, si G et G' sont isomorphes, alors le groupe G est divisible si, et seulement si, G' l'est.*

(3) *Tout quotient d'un groupe divisible est divisible.*

(4) *Soit H un sous-groupe de G. Si H et G/H sont divisibles, il en est de même de G.*

(5) *Si $G = \prod_{i\in I} G_i$, alors G est divisible si, et seulement si, G_i est divisible pour tout $i \in I$.*

(6) *Si $G = \coprod_{i\in I} G_i$, alors G est divisible si, et seulement si, G_i est divisible pour tout $i \in I$.*

(7) *Si $G = \bigoplus_{i\in I} H_i$, alors G est divisible si, et seulement si, H_i est divisible pour tout $i \in I$.*

(8) *Si G est p-primaire, alors G est divisible si, et seulement si, $G = p\cdot G$.*

Démonstration

(1) Soit $n \geqslant 1$ un entier et soit $x \in \mathrm{Tor}(G)$. Par hypothèse, il existe $y \in G$ tel que $x = n{\cdot}y$. Puisque $x \in \mathrm{Tor}(G)$, il existe $m \geqslant 1$ tel que $m{\cdot}x = 0$. Mais alors, $mn{\cdot}y = 0$ et ainsi $y \in \mathrm{Tor}(G)$. On a donc $\mathrm{Tor}(G) = n{\cdot}\mathrm{Tor}(G)$. Ainsi, le groupe $\mathrm{Tor}(G)$ est divisible.

(2) Soit $n \geqslant 1$ un entier, et soit $x' \in \mathrm{Im}(f)$. On a donc $x' = f(x)$ pour un certain $x \in G$. Par hypothèse, il existe $y \in G$ tel que $x = n{\cdot}y$. On a alors

$$f(x) = f(n{\cdot}y) = n{\cdot}f(y).$$

Comme $f(y) \in \mathrm{Im}(f)$, on a bien

$$\mathrm{Im}(f) = n{\cdot}\big(\mathrm{Im}(f)\big).$$

Ainsi, $\mathrm{Im}(f)$ est divisible. Si maintenant $f : G \xrightarrow{\sim} G'$ est un isomorphisme, alors f et f^{-1} sont surjectives. Il suffit alors d'appliquer la première partie pour conclure.

(3) Il suffit d'appliquer (2) au morphisme surjectif $\pi : G \longrightarrow G/H$.

(4) Supposons que les groupes H et G/H soient divisibles. Soit $n \geqslant 1$ un entier et soit $x \in G$. Par hypothèse, il existe $\overline{y} \in G/H$ tel que

$$\overline{x} = n{\cdot}\overline{y} = \overline{n{\cdot}y}.$$

Ainsi, il existe $h \in H$ tel que $x = n{\cdot}y + h$. Comme le groupe H est divisible, il existe $h' \in H$ tel que $h = n{\cdot}h'$. On a alors $x = n{\cdot}(y+h')$, et donc $G = n{\cdot}G$. Par conséquent, G est divisible.

(5) Supposons que $G = \prod_{i \in I} G_i$ soit divisible. En appliquant le point (2) à la i-ème projection $G \longrightarrow G_i$, on obtient que G_i est divisible. Inversement, supposons que G_i soit divisible pour tout $i \in I$ et soit $n \geqslant 1$ un entier. Soit alors $x = (x_i)_{i \in I} \in G$. Par hypothèse, il existe $y_i \in G_i$ tel que $x_i = n{\cdot}y_i$. Posons $y = (y_i)_{i \in I}$. Ainsi, $x = n{\cdot}y$, et donc $G = n{\cdot}G$. Par suite, G est divisible si, et seulement si, tous les G_i le sont.

(6) La démonstration est identique à celle de (5), si ce n'est que pour tout indice $i \in I$ tel que $x_i = 0$, on pose $y_i = 0$, de façon à obtenir un élément y qui appartienne à $\coprod_{i \in I} G_i$.

(7) Puisque $G \simeq \coprod_{i \in I} H_i$, G est divisible si, et seulement si, $\coprod_{i \in I} H_i$ est divisible par le point (2). Cela équivaut au fait que H_i est divisible pour tout $i \in I$ d'après le point (6).

(8) Supposons que G soit un groupe p-primaire. Soit $n \geqslant 1$ un entier, et écrivons $n = p^r q$, avec $r \geqslant 0$ et $p \nmid q$. Soit $x \in G$. Par hypothèse, il existe un entier $m \geqslant 1$ tel que $p^m{\cdot}x = 0$.

Il existe alors u, $v \in \mathbb{Z}$ tels que $uq + vp^m = 1$. Mais alors,

$$x = (uq + vp^m)\cdot x = (uq)\cdot x = q\cdot(u\cdot x).$$

Cela dit démontre que $G = q\cdot G$. Mais, l'égalité $G = p\cdot G$ et une récurrence immédiate montre que l'on a $G = p^k\cdot G$ pour tout $k \geqslant 0$. Par conséquent, on a $G = q\cdot(p^r\cdot G) = n\cdot G$, et G est divisible. □

6.4. Remarque. Un sous-groupe d'un groupe abélien divisible n'est pas nécessairement divisible. Par exemple, $\mathbb{C}^\times$ est divisible, mais $\mathbb{R}^\times$ n'est pas divisible.

6.5. Exemple. Pour tout nombre premier p, le groupe

$$\mathbb{U}_{p^\infty} = \{z \in \mathbb{C} \mid \text{ il existe } n \geqslant 1, \text{ tel que } z^{p^n} = 1\}$$

est divisible. En effet, ce groupe est clairement p-primaire et vérifie

$$\mathbb{U}_{p^\infty} = \mathbb{U}_{p^\infty}^p.$$

On applique alors le lemme 6.3 (8) pour conclure.

Le but de ce qui suit est de décortiquer la structure des groupes abéliens divisibles. Le résultat suivant est crucial.

6.6. Théorème (Critère de Baer). *Soit G un groupe abélien. Alors, G est divisible si, et seulement si, pour tout groupe abélien A, pour tout sous-groupe B de A, et pour tout morphisme de groupes $g : B \longrightarrow G$, il existe un morphisme de groupes $f : A \longrightarrow G$ tel que $f_{|B} = g$.*

Démonstration. On suppose que G vérifie la propriété de l'énoncé.

Soit $n \geqslant 1$ et soit $x \in G$. Considérons l'application

$$\begin{aligned} g : n\mathbb{Z} &\longrightarrow G \\ nm &\longmapsto m\cdot x. \end{aligned}$$

Cette application est bien définie, car tout élément de $n\mathbb{Z}$ s'écrit de manière unique sous la forme nm, $m \in \mathbb{Z}$. On vérifie aisément que g est un morphisme de groupes.

Il existe donc un morphisme $f : \mathbb{Z} \longrightarrow G$ tel que

$$f(nm) = m\cdot x, \quad \text{pour tout } m \in \mathbb{Z}.$$

Posons $y = f(1)$. On a alors $f(n) = x = n\cdot f(1) = n\cdot y$, et donc $G = n\cdot G$. Cela étant vrai pour tout $n \geqslant 1$, le groupe G est divisible.

Réciproquement, soit G un groupe abélien divisible, soit A un groupe abélien, soit B un sous-groupe de A, et soit $g : B \longrightarrow G$ un morphisme de groupes.

Soit $\mathscr{F}$ l'ensemble des couples (H, θ), où H est un sous-groupe de A contenant B, et $\theta : H \longrightarrow G$ est un morphisme de groupes tel que $\theta_{|B} = g$.

On munit $\mathscr{F}$ de la relation « $\preceq$ » suivante : si $(H, \theta), (H', \theta') \in \mathscr{F}$, on dit que $(H, \theta) \preceq (H', \theta')$ si H est un sous-groupe de H' et $\theta'_{|H} = \theta$.

Nous allons démontrer que $(\mathscr{F}, \preceq)$ est inductif et non vide.

L'ensemble $\mathscr{F}$ contient (B, g) et est donc non vide. De plus, on vérifie aisément que la relation $\preceq$ est une relation d'ordre sur $\mathscr{F}$. Soit $(I, \leqslant)$ un ensemble totalement ordonné, et soit $(H_i, \theta_i)_{i \in I}$ une famille d'éléments de $\mathscr{F}$ totalement ordonnée par I, i.e. telle que, pour tous $i, j \in I$, on a

$$i \leqslant j \implies (H_i, \theta_i) \preceq (H_j, \theta_j).$$

On pose $H = \bigcup_{i \in I} H_i$. Si $x \in H$, on pose

$$\theta(x) = \theta_i(x),$$

où $i \in I$ est tel que $x \in H_i$. Montrons que $(H, \theta) \in \mathscr{F}$.

Puisque H contient B par construction, H est non vide. Soient $x, y \in H$. Alors, il existe $i, j \in I$ tels que $x \in H_i$ et $y \in H_j$. Si $k = \max(i, j)$, on a donc $x, y \in H_k$ par hypothèse sur les H_i. Comme H_k est un sous-groupe de A, on a $x - y \in H_k \subset H$. Par conséquent, H est un sous-groupe de A contenant B.

Montrons que $\theta : H \longrightarrow G$ est bien définie. Soit $x \in H$. Si $i, j \in I$ sont deux indices tels que $x \in H_i \cap H_j$, alors $x \in H_k$, où $k = \max(i, j)$. Puisque l'on a $i \leqslant k$ et $j \leqslant k$, on obtient $(H_i, \theta_i) \preceq (H_k, \theta_k)$ et $(H_j, \theta_j) \preceq (H_k, \theta_k)$. On a alors

$$\theta_k(x) = \theta_i(x) \quad \text{et} \quad \theta_k(x) = \theta_j(x).$$

Ainsi, $\theta_i(x) = \theta_j(x)$, et $\theta(x)$ ne dépend pas de l'indice $i \in I$ choisi.

Montrons que θ est un morphisme de groupes. Si $x, y \in H$, il existe $i, j \in I$ tels que $x \in H_i$ et $y \in H_j$. Si $k = \max(i, j)$, on a donc $x, y \in H_k$, et ainsi

$$\theta(x + y) = \theta_k(x + y) = \theta_k(x) + \theta_k(y) = \theta(x) + \theta(y).$$

Enfin, $\theta_{|B} = g$ puisque tous les morphismes θ_i prolongent g. On a donc montré que $(H, \theta) \in \mathscr{F}$. De plus, par construction, pour tout $i \in I$, H_i est un sous-groupe de H, et $\theta_{|H_i} = \theta_i$. Par conséquent, on a

$$(H_i, \theta_i) \preceq (H, \theta), \quad \text{pour tout } i \in I,$$

et $\mathscr{F}$ est donc un ensemble inductif non vide. Par le lemme de Zorn, $\mathscr{F}$ possède un élément maximal (H_0, θ_0).

On va maintenant montrer par l'absurde que $H_0 = A$. Supposons au contraire que $H_0 \neq A$, et soit $x \in A \setminus H_0$. Soit $H_1 = \langle H_0, x\rangle$.

Supposons tout d'abord que $H_0 \cap \langle x\rangle = \{0\}$. Dans ce cas, $H_1 = H_0 \oplus \langle x\rangle$, d'après la proposition X-2.5, par exemple. L'application

$$\begin{array}{rccc} \theta_1 : & H_1 & \longrightarrow & G \\ & h + m{\cdot}x & \longmapsto & \theta_0(h) \end{array}$$

est alors un morphisme de groupes vérifiant $(\theta_1)_{|_{H_0}} = \theta_0$. En particulier, on a $(\theta_1)_{|_B} = g$, et donc $(H_1, \theta_1) \in \mathscr{F}$. Comme H_0 est un sous-groupe strict de H_1, on a donc

$$(H_0, \theta_0) \prec (H_1, \theta_1),$$

ce qui contredit la maximalité de (H_0, θ_0).

On a donc $H_0 \cap \langle x\rangle \neq \{0\}$. Remarquons que $\overline{x}$ est d'ordre fini dans A/H_0. En effet, par hypothèse, il existe $m \in \mathbb{Z}$ non nul tel que $m{\cdot}x \in H_0$. Mais alors

$$m{\cdot}\overline{x} = \overline{m{\cdot}x} = \overline{0} \in A/H_0.$$

Soit $d \geqslant 1$ l'ordre de $\overline{x} \in A/H_0$.

Tout élément $z \in H_1$ s'écrit $z = h + m{\cdot}x$, avec $h \in H_0$ et $m \in \mathbb{Z}$. En effectuant la division euclidienne de m par d, on voit que z s'écrit

$$z = h + r{\cdot}x, \quad h \in H_0,\ r \in [\![0, d-1]\!].$$

Cette écriture est unique. En effet, supposons que $z = h + r{\cdot}x = h' + r'{\cdot}x$, avec $h,\ h' \in H_0$ et $r,\ r' \in [\![0, d-1]\!]$. On peut toujours supposer que $r \leqslant r'$. Mais alors, on a $(r' - r){\cdot}x = h - h' \in H_0$, et donc $(r' - r){\cdot}\overline{x} = \overline{0} \in A/H_0$. Ainsi, $d \mid (r' - r)$. Puisque $r' - r \in [\![-(d-1), d-1]\!]$, on a donc $r' - r = 0$, soit encore $r' = r$. Mais alors, $h' = h$, d'où la conclusion voulue.

Puisque G est divisible, il existe donc $y \in G$ tel que $d{\cdot}y = \theta_0(d{\cdot}x)$. Pour tout $r \in [\![0, d-1]\!]$, et tout $h \in H_0$, on pose

$$\theta_1(h + r{\cdot}x) = \theta_0(h) + r{\cdot}y.$$

Nous allons montrer que $\theta_1 : H_1 \longrightarrow G$ est un morphisme de groupes. Soient $h + r{\cdot}x,\ h' + r'{\cdot}x \in H_1$, avec $h,\ h' \in H_0$ et $r,\ r' \in [\![0, d-1]\!]$.

Si $r + r' \in [\![0, d-1]\!]$, on a

$$\begin{array}{rcl} \theta_1\big((h + r{\cdot}x) + (h' + r'{\cdot}x)\big) & = & \theta_1\big((h + h') + (r + r'){\cdot}x\big) \\ & = & \theta_0(h + h') + (r + r'){\cdot}y \\ & = & \theta_0(h) + \theta_0(h') + r{\cdot}y + r'{\cdot}y \\ & = & \big(\theta_0(h) + r{\cdot}y\big) + \big(\theta_0(h') + r'{\cdot}y\big) \\ & = & \theta_1(h + r{\cdot}x) + \theta_1(h' + r'{\cdot}x). \end{array}$$

Si $r + r' \in [\![d, 2d-2]\!]$, alors $r + r' - d \in [\![0, d-1]\!]$, et l'on a

$$\begin{aligned}
\theta_1\big((h + r\cdot x) + (h' + r'\cdot x)\big) &= \theta_1\big((h + h' + d\cdot x) + (r + r' - d)\cdot x\big) \\
&= \theta_0(h + h' + d\cdot x) + (r + r' - d)\cdot y \\
&= \theta_0(h) + \theta_0(h') + \theta_0(d\cdot x) + r\cdot y + r'\cdot y - dy \\
&= \theta_0(h) + \theta_0(h') + d\cdot y + r\cdot y + r'\cdot y - d\cdot y \\
&= \big(\theta_0(h) + r\cdot y\big) + \big(\theta_0(h') + r'\cdot y\big) \\
&= \theta_1(h + r\cdot x) + \theta_1(h' + r'\cdot x).
\end{aligned}$$

Ainsi, θ_1 est un morphisme de groupes tel que $(\theta_1)_{|B} = g$, et $(H_1, \theta_1) \in \mathscr{F}$. De plus, on a

$$(H_0, \theta_0) \prec (H_1, \theta_1).$$

Cela contredit une nouvelle fois la maximalité de (H_0, θ_0).

Finalement, on a $H_0 = A$ et $f = \theta_0$ est le morphisme recherché. Cela achève la démonstration. □

Comme $\mathbb{C}^\times$ est un groupe abélien divisible, on obtient immédiatement le résultat suivant [3], qui généralise le théorème 3.4.

6.7. Corollaire. *Soit G un groupe abélien, et soit H un sous-groupe. Alors, le morphisme de restriction*

$$\theta : \widehat{G} \longrightarrow \widehat{H}$$

est surjectif.

Nous allons maintenant effectuer un premier dévissage des groupes abéliens divisibles.

6.8. Proposition. *Soit $n \geqslant 1$ et soit G un groupe abélien. Alors, G est divisible si, et seulement si, il existe un sous-groupe D et une famille de sous-groupes $(H_p)_p$ vérifiant les conditions suivantes :*

(1) *le groupe D est divisible sans torsion ;*

(2) *pour tout p premier, le sous-groupe H_p est divisible p-primaire ;*

(3) *on a $G = D \oplus \bigoplus_p H_p$.*

Dans ce cas, le sous-groupe H_p est la composante p-primaire de G pour tout p, et l'on a $D \simeq G/\mathrm{Tor}(G)$. En particulier, D et H_p (p premier) sont uniques à isomorphisme près.

Démonstration. Supposons que G soit un groupe divisible. Alors, $\mathrm{Tor}(G)$ est divisible par le lemme 6.3 (1). De plus, $\mathrm{Tor}\big(G/\mathrm{Tor}(G)\big) = \{0\}$ d'après le lemme 1.4. Par le théorème 6.6 appliqué à $B = \mathrm{Tor}(G)$, avec $A = G$ et $g = \mathrm{Id}_{\mathrm{Tor}(G)}$, il existe un morphisme de groupes $\pi : G \longrightarrow \mathrm{Tor}(G)$ tel que $\pi_{|\mathrm{Tor}(G)} = \mathrm{Id}_{\mathrm{Tor}(G)}$.

3. Attention au conflit de notation !

Si $D = \mathrm{Ker}(\pi)$, on a donc $G = D \oplus \mathrm{Tor}(G)$ par le corollaire X-3.18. Ainsi, on a que $D \simeq G/\mathrm{Tor}(G)$, et D est en particulier sans torsion. Par le théorème 1.8, $\mathrm{Tor}(G)$ est somme directe interne des composantes p-primaires G_p de G. On obtient donc

$$G = D \oplus \bigoplus_p G_p.$$

Comme G est divisible, par le lemme 6.3 (7), D et G_p sont divisibles.

Inversement, supposons que G possède une décomposition satisfaisant les trois conditions de l'énoncé. Si D est divisible, et si H_p est divisible pour tout p, alors G est divisible par le lemme 6.3 (7).

Soit p un nombre premier. Comme H_p est p-primaire, on a donc $H_p \subset G_p$ par définition. Inversement, soit $x \in G_p$. Il existe donc un entier $m \geqslant 1$ tel que $p^m \cdot x = 0$. Écrivons

$$x = x_D + \sum_\ell x_\ell, \quad x_D \in D,\ x_\ell \in H_\ell.$$

On a donc $0 = p^m \cdot x_D + \sum_\ell p^m \cdot x_\ell$. Les groupes D et H_ℓ (ℓ premier) étant en somme directe, on obtient

$$p^m \cdot x_D = 0 \quad \text{et} \quad p^m \cdot x_\ell = 0, \quad \text{pour tout } \ell \text{ premier.}$$

Puisque D est sans torsion, on en déduit $x_D = 0$. Soit $\ell \neq p$. Par hypothèse sur H_ℓ, il existe $n \geqslant 1$ tel que $\ell^n \cdot x_\ell = 0$. Ainsi, l'ordre de x_ℓ est à la fois une puissance de p et et une puissance de ℓ. Par conséquent, $x_\ell = 0$. Finalement, on en déduit $x = x_p \in H_p$, et $H_p = G_p$. Mais alors, on a

$$\bigoplus_p H_p = \bigoplus_p G_p = \mathrm{Tor}(G),$$

d'après le théorème 1.8. On a donc $G = D \oplus \mathrm{Tor}(G)$, et par conséquent, on a $D \simeq G/\mathrm{Tor}(G)$. Le dernier point est alors immédiat, et cela achève la démonstration. □

Nous allons maintenant nous intéresser aux groupes divisibles p-primaires.

Notation. Si G est un groupe abélien, on note

$$G[p] = \{x \in G \mid p \cdot x = 0\}.$$

On montre aisément que l'application

$$\begin{aligned} \mathbb{F}_p \times G[p] &\longrightarrow G[p] \\ (\overline{m}, x) &\longmapsto \overline{m \cdot x} \end{aligned}$$

est bien définie, et induit sur $G[p]$ une structure de $\mathbb{F}_p$-espace vectoriel.

6.9. Remarque. Tout isomorphisme de groupes $G \simeq G'$ induit un isomorphisme $G[p] \simeq G'[p]$ pour tout p premier.

En effet, soit $f : G \xrightarrow{\sim} G'$ un isomorphisme. Pour tout $x \in G$, on a

$$p{\cdot}x = 0 \iff f(p{\cdot}x) = 0 \iff p{\cdot}f(x) = 0.$$

On a donc $x \in G[p]$ si, et seulement si, $f(x) \in G'[p]$. Ainsi, f induit un isomorphisme

$$G[p] \simeq G'[p].$$

6.10. Lemme. *Soient G, G' deux groupes divisibles p-primaires. Alors, G et G' sont isomorphes si, et seulement si, $G[p]$ et $G'[p]$ sont isomorphes.*

Démonstration. Le sens direct est donné par la remarque 6.9.
Inversement, soit $h : G[p] \xrightarrow{\sim} G'[p]$ un isomorphisme. En composant par l'inclusion $G'[p] \subset G'$, on obtient un morphisme $G[p] \longrightarrow G'$, qui s'étend en un morphisme $f : G \longrightarrow G'$ par le théorème 6.6. Par définition de f, on a alors

$$f(z) = h(z), \quad \text{pour tout } z \in G[p].$$

Montrons que f est un isomorphisme.

Soit $x \in \mathrm{Ker}(f)$. Comme G est p-primaire, il existe $m \geqslant 1$ tel que $p^m{\cdot}x = 0$. On a donc $p^{m-1}{\cdot}x \in G[p]$. Mais alors

$$h(p^{m-1}{\cdot}x) = f(p^{m-1}{\cdot}x) = p^{m-1}{\cdot}f(x) = 0.$$

Comme h est un isomorphisme de groupes, on obtient $p^{m-1}{\cdot}x = 0$. Par récurrence sur m, on obtient finalement que $x = 0$, et f est donc injective.

Montrons maintenant la surjectivité de f. Il suffit pour cela de vérifier la propriété suivante pour tout $m \geqslant 0$:

(H_m) Tout élément $y \in G'$ vérifiant $p^m{\cdot}y = 0$ est dans l'image de f.

Nous allons procéder par récurrence. C'est clairement vrai pour $m = 0$. Supposons que (H_m) soit vraie pour un $m \geqslant 0$, et montrons que (H_{m+1}) est vraie. Soit $y \in G'$ vérifiant $p^{m+1}{\cdot}y = 0$. Alors, $p^m{\cdot}(p{\cdot}y) = 0$.

Par hypothèse de récurrence, $p{\cdot}y \in \mathrm{Im}(f)$. Puisque le groupe G est divisible, $\mathrm{Im}(f)$ est divisible, par le lemme 6.3 (2). Il existe donc $x \in G$ tel que $p{\cdot}y = p{\cdot}f(x)$. Mais alors, on a $y - f(x) \in G'[p]$. Comme h est un isomorphisme, il existe $x' \in G[p]$ tel que $y - f(x) = h(x') = f(x')$.
Ainsi, $y = f(x+x') \in \mathrm{Im}(f)$, ce qui achève la récurrence. Par conséquent, f est bien surjective, et f est l'isomorphisme recherché. □

On peut maintenant démontrer la proposition suivante.

6.11. Proposition. *Soit G un groupe abélien divisible p-primaire. Alors, il existe un ensemble I tel que*

$$G \simeq \mathbb{U}_{p^\infty}^{(I)},$$

où $\mathbb{U}_{p^\infty} = \{z \in \mathbb{C} \mid$ *il existe* $n \geqslant 1$ *tel que* $z^{p^n} = 1\}$.

Démonstration. D'après l'exemple 6.5, $\mathbb{U}_{p^\infty}$ est divisible p-primaire. Enfin, on a aisément

$$\mathbb{U}_{p^\infty}[p] = \mathbb{U}_p \simeq \mathbb{Z}/p\mathbb{Z}.$$

Comme $G[p]$ est un $\mathbb{F}_p$-espace vectoriel, il existe un ensemble I tel que

$$G[p] \simeq (\mathbb{Z}/p\mathbb{Z})^{(I)}.$$

En utilisant les définitions, on voit que l'on a

$$(\mathbb{U}_{p^\infty}^{(I)})[p] = (\mathbb{U}_{p^\infty}[p])^{(I)}.$$

Ainsi, on a un isomorphisme de groupes

$$G[p] \simeq (\mathbb{Z}/p\mathbb{Z})^{(I)} \simeq (\mathbb{U}_{p^\infty}[p])^{(I)} \simeq (\mathbb{U}_{p^\infty}^{(I)})[p],$$

d'où un isomorphisme $G \simeq \mathbb{U}_{p^\infty}^{(I)}$ d'après la proposition précédente. Cela achève la démonstration. □

Il nous reste maintenant à élucider la structure des groupes abéliens divisibles sans torsion, ce qui est fait dans la proposition suivante.

6.12. Proposition. *Soit G un groupe abélien divisible sans torsion. Alors, G est un $\mathbb{Q}$-espace vectoriel. En particulier, il existe un ensemble I tel que*

$$G \simeq \mathbb{Q}^{(I)}.$$

Démonstration. Soit $x \in G$ et soit $r = \frac{a}{b} \in \mathbb{Q}$, avec $b \in \mathbb{Z} \setminus \{0\}$. Alors, les égalités $G = |b| \cdot G = b \cdot G$ montrent qu'il existe $y \in G$ tel que $x = b \cdot y$. De plus, un tel $y \in G$ est unique. En effet, si $x = b \cdot y_1 = b \cdot y_2$, avec $y_1, y_2 \in G$, on a $b \cdot (y_1 - y_2) = 0$. Mais alors, $y_1 - y_2 \in \mathrm{Tor}(G) = \{0\}$, d'où $y_1 = y_2$. Pour tout $r = \frac{a}{b} \in \mathbb{Q}$, on pose alors

$$r * x = a \cdot y,$$

où $y \in G$ est l'unique élément de G tel que $x = b \cdot y$.

Vérifions que $r * x$ ne dépend que de r. Si $r = \frac{a}{b} = \frac{c}{d}$, soit $z \in G$ l'unique élément de G tel que $x = d \cdot z$. On doit vérifier que $c \cdot z - a \cdot y = 0$. Comme G est sans torsion, cela équivaut à $bd \cdot (c \cdot z - a \cdot y) = 0$.

Mais, on a

$$bd\cdot(c\cdot z - a\cdot y) = bc\cdot(d\cdot z) - ad\cdot(b\cdot y) = bc\cdot x - ad\cdot x = (bc - ad)\cdot x = 0,$$

d'où le résultat voulu. Ainsi, l'application

$$\begin{aligned} \mathbb{Q} \times G &\longrightarrow G \\ (r, x) &\longmapsto r * x \end{aligned}$$

est bien définie. Il reste à vérifier que cette loi confère à $(G, +)$ une structure de $\mathbb{Q}$-espace vectoriel. Soient $x, x' \in G$, $\frac{a}{b}, \frac{a'}{b'} \in \mathbb{Q}$. Il est clair que $1 * x = x$. Soient $y, y' \in G$ vérifiant $x = b\cdot y$ et $x' = b\cdot y'$. Alors, $x + x' = b\cdot(y + y')$. On obtient donc

$$\frac{a}{b} * (x + x') = a\cdot(y + y') = a\cdot y + a\cdot y' = \frac{a}{b} * x + \frac{a}{b} * x'.$$

Soient $z, t \in G$ vérifiant $x = b'\cdot z$ et $x = bb'\cdot t$. Comme $x = b\cdot y$, on a $y = b'\cdot t$ par unicité de y. De même, on a $z = b\cdot t$. On a alors

$$\begin{aligned} \left(\frac{a}{b} + \frac{a'}{b'}\right) * x &= \frac{ab' + a'b}{bb'} * x \\ &= (ab' + a'b)\cdot t \\ &= a\cdot(b'\cdot t) + a'\cdot(b\cdot t) \\ &= a\cdot y + a'\cdot z \\ &= \frac{a}{b} * x + \frac{a'}{b'} * x. \end{aligned}$$

Enfin, en gardant les mêmes notations, on a

$$\left(\frac{a}{b}\,\frac{a'}{b'}\right) * x = \frac{aa'}{bb'} * x = (aa')\cdot t = a\cdot(a'\cdot t).$$

Mais on a aussi $\frac{a'}{b'} * x = a'\cdot z$. Comme $a'\cdot z = a'b\cdot t = b\cdot(a'\cdot t)$, on a alors

$$\left(\frac{a}{b}\,\frac{a'}{b'}\right) * x = \frac{a}{b} * (a'\cdot z) = \frac{a}{b} * \left(\frac{a'}{b'} * x\right).$$

Ainsi, G est bien un $\mathbb{Q}$-espace vectoriel, et le choix d'une base $\mathbf{e} = (e_i)_{\in I}$ fournit alors un isomorphisme de $\mathbb{Q}$-espaces vectoriels

$$G \simeq \mathbb{Q}^{(I)},$$

qui est en particulier un isomorphisme de groupes. □

En appliquant les propositions 6.8, 6.12 et 6.11, on obtient immédiatement le théorème de structure suivant.

6.13. Théorème. *Soit G un groupe abélien divisible. Alors, il existe un ensemble I, et, pour tout p premier, un ensemble I_p, tels que*

$$G \simeq \mathbb{Q}^{(I)} \times \coprod_p \mathbb{U}_{p^\infty}^{(I_p)}.$$

On complète ce théorème de structure par un résultat de classification.

6.14. Théorème. *Deux groupes abéliens divisibles G et G' sont isomorphes si, et seulement si, les conditions suivantes sont vérifiées :*

(1) *les $\mathbb{Q}$-espaces vectoriels $G/\mathrm{Tor}(G)$ et $G'/\mathrm{Tor}(G')$ sont isomorphes ;*

(2) *pour tout p premier, les $\mathbb{F}_p$-espaces vectoriels $G[p]$ et $G'[p]$ sont isomorphes.*

Démonstration. Soient maintenant G et G' deux groupes abéliens divisibles. Si $G \simeq G'$, alors on a déjà vu que $\mathrm{Tor}(G) \simeq \mathrm{Tor}(G')$. Mais alors, l'isomorphisme $G \simeq G'$ se factorise en un isomorphisme

$$G/\mathrm{Tor}(G) \simeq G'/\mathrm{Tor}(G'),$$

et l'on vérifie sans mal que c'est aussi un isomorphisme de $\mathbb{Q}$-espaces vectoriels. D'après la remarque 6.9, on a aussi $G[p] \simeq G'[p]$. On vérifie que c'est également un isomorphisme de $\mathbb{F}_p$-espaces vectoriels.

Inversement, soient G et G' deux groupes abéliens divisibles vérifiant les conditions de l'énoncé. Par la proposition 6.8, on peut écrire

$$G = D \oplus \bigoplus_p H_p \quad \text{et} \quad G' = D' \oplus \bigoplus_p H'_p,$$

où D et D' sont des sous-groupes divisibles sans torsion, et où H_p et H'_p sont des sous-groupes divisibles p-primaires. Toujours d'après cette même proposition, on a $D \simeq G/\mathrm{Tor}(G)$, $D' \simeq G'/\mathrm{Tor}(G')$, $H_p = G_p$ et $H'_p = G'_p$. Par hypothèse, on a donc $D \simeq D'$.

De plus, le sous-groupe $G[p]$ des éléments d'ordre divisant p est inclus dans G_p, et l'on a clairement les égalités

$$G[p] = G_p[p] = H_p[p].$$

De même, on a $G'[p] = H'_p[p]$.

Par hypothèse, on a alors un isomorphisme de $\mathbb{F}_p$-espaces vectoriels

$$H_p[p] \simeq H'_p[p],$$

et le lemme 6.10 montre que $H_p \simeq H'_p$. On déduit de tout cela que l'on a un isomorphisme $G \simeq G'$. Cela achève la démonstration. □

6.15. Remarque. Soit G un groupe abélien divisible, et écrivons

$$G \simeq \mathbb{Q}^{(I)} \times \coprod_p \mathbb{U}_{p^\infty}^{(I_p)}.$$

On vérifie aisément que $G/\mathrm{Tor}(G) \simeq \mathbb{Q}^{(I)}$, et que pour tout p, on a un

isomorphisme de $\mathbb{F}_p$-espaces vectoriels

$$G[p] \simeq \mathbb{F}_p^{(I_p)}.$$

Or, on peut démontrer que, dans un espace vectoriel quelconque, toutes les bases sont équipotentes (cf. [9, Théorème 2.11], par exemple). On déduit alors du théorème précédent que les ensembles I et I_p, avec p premier, sont uniques à équipotence près.

Les détails sont laissés au lecteur.

7. Exercices

Les quatre exercices suivants ne nécessitent que le contenu du paragraphe §2 de ce chapitre, ainsi que le théorème de structure des groupes abéliens finis et la définition des invariants de similitude. Ils peuvent ainsi être abordés par un étudiant de niveau L3.

E1. Soit G un groupe fini. Montrer que $\exp(G) = |G|$ si, et seulement si, tous les sous-groupes de Sylow de G sont cycliques.

E2.

a. Calculer les invariants de similitude de $G = \mathbb{Z}/45\mathbb{Z} \times \mathbb{Z}/15\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z}$.

b. Calculer les invariants de similitude de $G = \mathbb{Z}/12\mathbb{Z} \times \mathbb{Z}/10\mathbb{Z} \times \mathbb{Z}/15\mathbb{Z}$.

E3. Trouver tous les groupes abéliens d'ordre 72, à isomorphisme près.

E4. Soit G un groupe abélien fini et soit p un nombre premier.

a. Soit $m \geqslant 1$. Montrer que l'application

$$\begin{array}{ccccc} \mathbb{F}_p & \times & \left(p^{m-1}\cdot G/p^m\cdot G\right) & \longrightarrow & p^{m-1}\cdot G/p^m\cdot G \\ (\tilde{m}, \overline{x}) & & & \longmapsto & \overline{m\cdot x} \end{array}$$

est bien définie.

Montrer ensuite qu'elle munit le groupe abélien $(p^{m-1}\cdot G/p^m\cdot G, +)$ d'une structure de $\mathbb{F}_p$-espace vectoriel de dimension finie.

b. Montrer que si $p \nmid |G|$, alors $p^{m-1}\cdot G/p^m\cdot G = \{\overline{0}\}$, pour tout $m \geqslant 1$.

c. Montrer que si $p \mid |G|$, alors pour tout $m \geqslant 0$, l'entier

$$\dim_{\mathbb{F}_p}\left(p^{m-1}\cdot G/p^m\cdot G\right) - \dim_{\mathbb{F}_p}\left(p^m\cdot G/p^{m+1}\cdot G\right)$$

est le nombre de facteurs isomorphes à $\mathbb{Z}/p^m\mathbb{Z}$ dans la décomposition de G en produit direct de groupes cycliques dont l'ordre est une puissance d'un nombre premier.

d. En déduire une autre démonstration de l'unicité de la suite des invariants de similitude d'un groupe abélien fini.

Les deux prochains exercices nécessitent d'avoir étudié le paragraphe IX-3.

E5. Soit G un groupe fini. Montrer que les propriétés suivantes sont équivalentes :

(1) le groupe G est résoluble ;

(2) il existe une suite de composition $(G_0, \ldots, G_r)$ de G telle que, pour tout $i \in [\![0, r-1]\!]$, le groupe G_i/G_{i+1} est cyclique d'ordre premier ;

(3) il existe une suite de composition $(G_0, \ldots, G_r)$ de G telle que, pour tout $i \in [\![0, r-1]\!]$, le groupe G_i/G_{i+1} est cyclique.

E6. Soit G un groupe abélien fini. Quels sont les quotients successifs d'une suite de Jordan-Hölder de G ?

Les exercices ci-dessous nécessitent d'avoir lu les paragraphes §1 – §5 de ce chapitre.

E7.

a. Soient G_1, G_2 des groupes abéliens.

(i) Montrer que si $G_1 \simeq G_2$, alors $\widehat{G}_1 \simeq \widehat{G}_2$.

(ii) Montrer que le groupe des caractères linéaires de $G_1 \times G_2$ est isomorphe à $\widehat{G}_1 \times \widehat{G}_2$.

b. En déduire que, pour tout groupe abélien fini G, on a un isomorphisme de groupes $G \simeq \widehat{G}$.

E8. Soit G un groupe abélien, et soit H un sous-groupe de G tel que G/H est sans torsion. Montrer que $\mathrm{Tor}(G) \subset H$.

E9. Vrai ou faux ? Soit L un groupe abélien libre de rang n. Alors :

(1) on a $L \simeq \langle a_1, \ldots, a_n \mid a_i a_j = a_j a_i$, pour tout $i < j \rangle$;

(2) toute famille libre de L peut être complétée en une base ;

(3) de toute famille génératrice de L, on peut en extraire une base ;

(4) toute famille libre de L à n éléments est une base ;

(5) tout groupe abélien engendré par n éléments est isomorphe à un quotient de L ;

(6) le groupe des automorphismes de L est isomorphe à

$$\mathrm{GL}_n(\mathbb{Z}) = \{M \in \mathrm{M}_n(\mathbb{Z}) \mid \det(M) = \pm 1\}.$$

E10. Montrer que pour tous groupes abéliens de type fini G, G_1, G_2, on a

$$G_1 \times G \simeq G_2 \times G \implies G_1 \simeq G_2.$$

Les exercices ci-après nécessitent d'avoir étudié le paragraphe §6 de ce chapitre.

E11.

a. Montrer que $\mathbb{R}/\mathbb{Z} \simeq \mathbb{R} \times \mathbb{Q}/\mathbb{Z}$.

Indication. On pourra remarquer que $\mathbb{R}/\mathbb{Z}$ est divisible.

b. Montrer que $\mathbb{Q}/\mathbb{Z} \simeq \coprod_p \mathbb{U}_{p^\infty}$.

E12. Montrer qu'un groupe abélien divisible de type fini est nul.

Indication. Procéder par récurrence sur le cardinal d'une famille génératrice.

E13. Soit $G = \prod_p \mathbb{Z}/p\mathbb{Z}$, où p décrit l'ensemble des nombres premiers.

a. Montrer que $\mathrm{Tor}(G) = \coprod_p \mathbb{Z}/p\mathbb{Z}$.

b. Montrer que $G/\mathrm{Tor}(G)$ est divisible.

c. Montrer que par l'absurde que $\mathrm{Tor}(G)$ n'est pas un facteur direct de G, i.e. qu'il n'existe pas de sous-groupe L de G tel que $G = L \oplus \mathrm{Tor}(G)$.

Indication. Montrer que si $G = L \oplus \mathrm{Tor}(G)$, alors L est divisible. En déduire que pour tout $x \in G$, il existe $d \geqslant 1$ tel que $d{\cdot}x \in n{\cdot}G$ pour tout $n \geqslant 1$. Considérer $x = (\overline{1}, \overline{1}, \ldots)$ et conclure.

E14. Montrer qu'un groupe abélien G est divisible si, et seulement si, pour tout groupe abélien G' contenant G, il existe un sous-groupe H' de G' tel que $G' = G \oplus H'$.

E15. Soit G un groupe abélien. On note $\Delta(G)$ le sous-groupe de G engendré par tous les sous-groupes divisibles de G. On dit que le groupe G est réduit si $\Delta(G) = \{0\}$.

a. Montrer que $\Delta(G)$ est le plus grand sous-groupe divisible de G.

b. Montrer qu'il existe un sous-groupe réduit R tel que $G = \Delta(G) \oplus R$.

c. Soient G, G' deux groupes abéliens. Montrer que $G \simeq G'$ si, et seulement si, $\Delta(G) \simeq \Delta(G')$ et $G/\Delta(G) \simeq G'/\Delta(G')$.

E16. Soit G un groupe abélien.

a. Montrer qu'il existe un morphisme injectif $G \hookrightarrow D$, où D est divisible.

Indication. Commencer par le cas où G est un groupe abélien libre. Montrer que dans ce cas, G s'identifie à un sous-groupe d'un $\mathbb{Q}$-espace vectoriel.

b. On suppose que G est sans torsion. Déduire de la question précédente qu'il existe un morphisme injectif $G \hookrightarrow E$, où E est un $\mathbb{Q}$-espace vectoriel.

E17. En admettant que toutes les bases d'un espace vectoriel sont équipotentes, démontrer les résultats de la remarque XII-6.15.

E18. Soit G un groupe abélien. Montrer que les conditions suivantes sont équivalentes :

(1) le groupe G est divisible ;

(2) tout quotient non nul de G est infini ;

(3) le groupe G ne possède pas de sous-groupe maximal.

Chapitre XIII

Quand tout groupe fini d'ordre n est-il cyclique ?

Si $n \geqslant 1$ est un entier, il existe au moins un groupe d'ordre n, à savoir le groupe $\mathbb{Z}/n\mathbb{Z}$. Il n'y a bien entendu aucune raison pour que $\mathbb{Z}/n\mathbb{Z}$ soit l'unique groupe d'ordre n à isomorphisme près, et c'est d'ailleurs faux en général, comme le montre déjà le cas des groupes d'ordre 4.

En revanche, cela peut arriver pour certaines valeurs de n. Par exemple, c'est vrai lorsque n est premier, mais aussi lorsque n est composé, comme par exemple pour $n = 15, 33, 35, 51$ (cf. exercice XI-4, **c**) ou $n = 255$ (cf. exercice XI-5).

Le but de ce qui suit est déterminer une condition nécessaire et suffisante sur $n \geqslant 1$ pour que tout groupe d'ordre n soit cyclique.

Ce chapitre ayant pour seul objectif de proposer une application des notions introduites jusqu'à présent, il ne contiendra aucun exercice.

1. Une condition nécessaire

Avant de nous lancer dans le vif du sujet, nous allons commencer modestement, et traiter le cas plus facile des groupes abéliens.

1.1. Proposition. *Soit $n \geqslant 1$ un entier. Alors, tout groupe abélien d'ordre n est cyclique si, et seulement si, n est produit de nombres premiers deux à deux distincts.*

Démonstration. Soit $n \geqslant 1$ un entier. Supposons tout d'abord que tout groupe abélien d'ordre n est cyclique.

Si $n = 1$, la condition de la proposition est clairement vérifiée. Supposons maintenant que $n \geqslant 2$, et qu'il existe un nombre premier p tel que $p^2 \mid n$. Écrivons $n = p^2 m$. Alors, $G = \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/pm\mathbb{Z}$ est un groupe abélien d'ordre n qui n'est pas cyclique, puisque tout élément de G est d'ordre divisant pm. Ainsi, n est un produit de nombres premiers deux à deux distincts.

Inversement, supposons que $n = p_1 \times \cdots \times p_r$, où $p_1, \ldots, p_r$ sont des nombres premiers deux à deux distincts, et soit G un groupe abélien d'ordre n. D'après le lemme de Cauchy, G possède un élément g_i d'ordre p_i. Comme les éléments $g_1, \ldots, g_r$ commutent deux à deux et ont des ordres premiers entre eux deux à deux, on a $o(g_1 \cdots g_r) = p_1 \times \cdots \times p_r$. Ainsi, G est cyclique, engendré par $g_1 \cdots g_r$. □

Augmentons un peu la difficulté, et traitons le cas des groupes quelconques dont l'ordre est produit de deux nombres premiers distincts.

1.2. Lemme. *Soient p, q deux nombres premiers, avec $p \leqslant q$. Alors, tout groupe d'ordre pq est cyclique si, et seulement si, $p < q$ et $q \not\equiv 1$ $[p]$.*

Démonstration. Si $p = q$, le groupe $(\mathbb{Z}/p\mathbb{Z})^2$ n'est pas cyclique.

Supposons $p < q$, et soit G un groupe d'ordre pq.

Le nombre N_q de p-Sylow de G vérifie $N_q \equiv 1$ $[q]$. Supposons que $N_q > 1$. Alors, $N_q \geqslant q + 1 > p$. Mais on a aussi, $N_q \mid p$. C'est donc 1 ou p, et les considérations précédentes donnent alors une contradiction. Par conséquent, $N_q = 1$, et un q-Sylow H de G est distingué dans G.

Soit K un p-Sylow de G. Comme les ordres de H et K sont q et p respectivement, ils sont premiers entre eux, et $H \cap K = \{1\}$. De plus, $|G| = |H||K|$. Par conséquent, $G = H \rtimes K \simeq H \times_\rho K$, où $\rho : K \longrightarrow \mathrm{Aut}(H)$ est un morphisme de groupes.

Comme H est cyclique d'ordre q premier, $\mathrm{Aut}(H)$ est isomorphe à $(\mathbb{Z}/q\mathbb{Z})^\times$, et est donc cyclique d'ordre $q - 1$ (cf. exercice VI-37 et corollaire 2.9).

Le morphisme ρ envoie un élément d'ordre p de K (qui existe, puisque K est cyclique) sur un élément de $\mathrm{Aut}(H)$ d'ordre divisant p.

Si $p \nmid q - 1$, $\mathrm{Aut}(H)$ n'a pas d'élément d'ordre p par Lagrange, et ρ est donc trivial. Cela signifie que $G \simeq H \times K$. Le lemme chinois donne alors

$$G \simeq \mathbb{Z}/q\mathbb{Z} \times \mathbb{Z}/p\mathbb{Z} \simeq \mathbb{Z}/pq\mathbb{Z},$$

ce qui signifie que G est cyclique.

Supposons maintenant $p \mid q - 1$. Nous allons construire un groupe fini d'ordre pq non cyclique. L'analyse précédente montre qu'il faut chercher G sous forme de produit semi-direct de $\mathbb{Z}/q\mathbb{Z}$ par $\mathbb{Z}/p\mathbb{Z}$.

Soit $\overline{a} \in (\mathbb{Z}/q\mathbb{Z})^\times$ un élément d'ordre p (qui existe, puisque $(\mathbb{Z}/q\mathbb{Z})^\times$ est cyclique d'ordre $q-1$ et $p \mid q-1$), et soit $\theta : \mathbb{Z}/q\mathbb{Z} \longrightarrow \mathbb{Z}/q\mathbb{Z}$ le morphisme de multiplication par $\overline{a}$. C'est un isomorphisme de groupes (son inverse est la multiplication par $\overline{a}^{-1}$).

Notons que θ^p est la multiplication par $\overline{a}^p = \overline{1}$, i.e. $\theta^p = \mathrm{Id}$. Mais alors, l'unique morphisme $\mathbb{Z} \longrightarrow \mathrm{Aut}(\mathbb{Z}/q\mathbb{Z})$ envoyant 1 sur θ se factorise en un morphisme $\rho_0 : \mathbb{Z}/p\mathbb{Z} \longrightarrow \mathrm{Aut}(\mathbb{Z}/q\mathbb{Z})$ non trivial. Le groupe $H \times_{\rho_0} K$ est alors un groupe d'ordre pq non abélien, donc en particulier non cyclique.

On a donc démontré que tout groupe d'ordre pq est cyclique si, et seulement si, $p < q$ et $p \nmid q-1$, ce qui est équivalent à l'énoncé du lemme. □

1.3. Remarque. Dans le cas où $p \mid q-1$, on peut en fait montrer que tous les produits semi-directs non triviaux de $\mathbb{Z}/p\mathbb{Z}$ par $\mathbb{Z}/q\mathbb{Z}$ sont tous isomorphes (cf. exercice XI-4).

Le lemme XIII-1.2 a déjà des conséquences intéressantes. Il permet en effet de dégager une condition nécessaire pour que tout groupe d'ordre n soit cyclique.

1.4. Lemme. *Soit $n \geqslant 1$ un entier. Si tout groupe d'ordre n est cyclique, alors $n = p_1 \times \cdots \times p_r$, où $r \geqslant 0$ et $p_1, \ldots, p_r$ sont des nombres premiers tels que $p_1 < \cdots < p_r$ et $p_j \not\equiv 1 \ [p_i]$ pour tous $1 \leqslant i < j \leqslant r$.*

Démonstration. Si tout groupe d'ordre n est cyclique, c'est en particulier le cas pour un groupe abélien. La proposition XIII-1.1 montre alors que n est un produit de nombres premiers deux à deux distincts.

Supposons maintenant que n possède au moins deux diviseurs premiers p, q vérifiant $p < q$ et $p \mid q-1$.

Écrivons $n = pqm$. D'après le lemme XIII-1.2 et sa démonstration, il existe un groupe G' d'ordre pq non abélien. Le groupe $G = G' \times \mathbb{Z}/m\mathbb{Z}$ est alors d'ordre n, non abélien, donc non cyclique.

Cela achève la démonstration. □

Nous allons démontrer dans la suite que cette condition nécessaire est aussi suffisante. Avant toute chose, nous aurons besoin de résultats de théorie des groupes.

2. Compléments de théorie des groupes

2.1. Définition. Soit G un groupe, et soit P une partie de G. Le *normalisateur* de P dans G est l'ensemble

$$N_G(P) = \{g \in G \mid gPg^{-1} = P\}.$$

C'est un sous-groupe de G contenant P.

Le *centralisateur* de P dans G est l'ensemble

$$C_G(P) = \{g \in G \mid gx = xg, \text{ pour tout } x \in P\}.$$

C'est un sous-groupe de $N_G(P)$ qui ne contient pas nécessairement P.

2.2. Remarque. Si H est un sous-groupe de G, on peut démontrer que le normalisateur $N_G(H)$ est le plus grand sous-groupe de G (au sens de l'inclusion) dans lequel H est distingué (cf. exercice VI-15).

2.3. Lemme. *Soit G un groupe fini, et soit H un sous-groupe de G. Alors, le nombre de sous-groupes de G conjugués à H est égal à $[G : N_G(H)]$.*

Démonstration. Soit E l'ensemble des sous-groupes de G conjugués à H. Par définition, G agit sur E par conjugaison, et cette action est transitive, puisque l'orbite de H est E. Par ailleurs, le stabilisateur de H sous cette action n'est rien d'autre que $N_G(H)$. Le cardinal de l'orbite de H, c'est-à-dire E, est donc $[G : N_G(H)]$. □

On finit par un lemme classique.

2.4. Lemme. *Soit G un groupe, et soit $Z(G)$ son centre. Si $G/Z(G)$ est monogène, alors G est abélien.*

Démonstration. Choisissons un générateur $\overline{a} \in G/Z(G)$ de $G/Z(G)$. Pour tout $g \in G$, il existe donc $n_g \in \mathbb{Z}$ tel que $\overline{g} = \overline{a}^{n_g}$. Il existe alors $z_g \in Z(G)$ tel que $g = a^{n_g} z_g$.

Soient maintenant $g_1, g_2 \in G$. Avec les notations précédentes, en utilisant le fait que $z_{g_1}, z_{g_2} \in Z(G)$ et que les puissances de a commutent entre elles, on a

$$g_1 g_2 = a^{n_{g_1}} z_{g_1} a^{n_{g_2}} z_{g_2} = a^{n_{g_1}} a^{n_{g_2}} z_{g_1} z_{g_2}$$

d'une part, et

$$g_2 g_1 = a^{n_{g_2}} z_{g_2} a^{n_{g_1}} z_{g_1} = a^{n_{g_2}} a^{n_{g_1}} z_{g_2} z_{g_1} = a^{n_{g_1}} a^{n_{g_2}} z_{g_1} z_{g_2}$$

d'autre part, soit $g_1 g_2 = g_2 g_1$. □

3. Le théorème principal

Nous allons maintenant démontrer que la condition nécessaire donnée par le lemme XIII-1.4 est aussi suffisante. Autrement dit, nous allons démontrer le théorème suivant.

3.1. Théorème. *Soit $n \geqslant 1$ un entier. Alors, tout groupe d'ordre n est cyclique si, et seulement si, $n = p_1 \times \cdots \times p_r$, où $p_1, \ldots, p_r$ sont des nombres premiers tels que $p_1 < \cdots < p_r$ et $p_j \not\equiv 1\ [p_i]$, pour tous $1 \leqslant i < j \leqslant r$.*

Démonstration. Il suffit de démontrer la condition suffisante. Nous allons procéder par l'absurde. On dit qu'un entier $n \geqslant 1$ vérifie la condition (C) si n est de la forme $n = p_1 \times \cdots \times p_r$, où $p_1, \ldots, p_r$ sont des nombres premiers tels que $p_1 < \cdots < p_r$ et $p_j \not\equiv 1\ [p_i]$, pour tous $1 \leqslant i < j \leqslant r$.

On remarque que si n vérifie la condition (C), alors tous ses diviseurs positifs la vérifient également.

On suppose qu'il existe un entier $n \geqslant 1$ vérifiant la condition (C) mais pour lequel il existe un groupe G d'ordre n non cyclique. On choisit n minimal pour cette propriété. En particulier, $n \geqslant 2$ et n'est pas un nombre premier. On a donc $r \geqslant 2$.

Fait 1. Tout sous-groupe propre de G est cyclique, et tout quotient de G par un sous-groupe distingué non trivial est cyclique.

En effet, leurs ordres sont des diviseurs stricts de n, et vérifient alors la condition (C). On conclut alors par minimalité de n.

Fait 2. Le centre de G est trivial.

Supposons le contraire. Alors, $G/Z(G)$ est cyclique (donc monogène) par le fait 1, et le lemme XIII-2.4 montre que G est abélien. Mais, alors G est cyclique d'après la proposition XIII-1.1, d'où une contradiction.

Notre but est maintenant d'étudier les sous-groupes maximaux de G (cf. définition XI-2.7). Rappelons que de tels sous-groupes maximaux existent, d'après le lemme XI-2.8.

Fait 3. Soit U un sous-groupe maximal de G. Alors, U n'est pas trivial.

En effet, comme on l'a vu, n possède au moins deux diviseurs premiers distincts p_1 et p_2. Si U était trivial, il serait contenu strictement dans le sous-groupe engendré par un élément d'ordre p_1. Un tel sous-groupe étant un sous-groupe strict de G, on aurait une contradiction avec la maximalité de U.

Fait 4. Soit U un sous-groupe maximal de G. Alors, pour tout $x \in U \backslash \{1_G\}$, on a $U = C_G(x)$.

Comme U est un sous-groupe propre de G, U est cyclique d'après le fait 1. Si $y \in U$ est un générateur de U, x est une puissance de y et commute donc avec x. Ainsi, $yxy^{-1} = x$. Par conséquent, $y \in C_G(x)$, et comme $C_G(x)$ est un sous-groupe de G, on a $\langle y \rangle = U \subset C_G(x)$. Or, $C_G(x)$ est un sous-groupe strict de G. Sinon, cela signifierait que $x \in Z(G)$, qui est trivial par le fait 2 et l'on aurait $x = 1_G$. Par conséquent, par maximalité de U, on obtient finalement $U = C_G(x)$.

Fait 5. Deux sous-groupes maximaux distincts s'intersectent trivialement.

Si U_1 et U_2 sont deux sous-groupes maximaux tels que $U_1 \cap U_2 \neq \{1_G\}$, choisissons $x \neq 1_G$ dans l'intersection. D'après le fait précédent, on a alors les égalités $U_1 = C_G(x) = U_2$.

Fait 6. Pour tout sous-groupe maximal U de G, on a $U = N_G(U)$.

C'est le point le plus délicat. Soit $y \in N_G(U)$. Par définition de $N_G(U)$, la conjugaison par y induit un automorphisme

$$\begin{array}{rccc} \alpha : & U & \longrightarrow & U \\ & x & \longmapsto & yxy^{-1}. \end{array}$$

Notons que α^n est la conjugaison par $y^n = 1_G$, donc l'ordre de α dans le groupe $\mathrm{Aut}(U)$ divise n. Or, U étant cyclique d'après le fait 1, $\mathrm{Aut}(U)$ est isomorphe à $(\mathbb{Z}/d\mathbb{Z})^\times$, où $d = |U|$. En particulier, l'ordre de α divise $\varphi(d)$. Comme $d \mid n$, on a $d = p_{i_1} \times \cdots \times p_{i_s}$, avec $1 \leqslant i_1 < i_2 < \cdots < i_s \leqslant r$, et par conséquent $\varphi(d) = (p_{i_1} - 1) \times \cdots \times (p_{i_s} - 1)$. Par hypothèse sur $p_1, \ldots, p_r$, aucun p_j ne divise un de ces facteurs. Ainsi, n et $\varphi(d)$ sont premiers entre eux et α est donc d'ordre 1. Autrement dit, $\alpha = \mathrm{Id}_U$ et y commute avec tout élément de U. Si $x \in U \setminus \{1_G\}$, on a alors $y \in C_G(x)$. Par le fait 4, on a donc $y \in U$. Autrement dit, $N_G(U) \subset U$, d'où l'égalité recherchée, puisque par définition, on a aussi $U \subset N_G(U)$.

Fait 7. Soit U un sous-groupe maximal de G. Alors, la réunion de tous les conjugués de U contient exactement $n - \dfrac{n}{|U|}$ éléments non triviaux.

D'après le lemme XIII-2.3, U possède $r = [G : N_G(U)]$ conjugués. Par le fait 6, on a $r = [G : U]$. Notons $g_1 U g_1^{-1}, \ldots, g_r U g_r^{-1}$ ces conjugués. Remarquons alors l'égalité

$$\Big(\bigcup_{i=1}^{r} g_i U g_i^{-1}\Big) \setminus \{1_G\} = \bigcup_{i=1}^{r} \big(g_i U g_i^{-1} \setminus \{1_G\}\big).$$

De plus, pour tout $i \in [\![1, r]\!]$, $g_i U g_i^{-1}$ est encore un sous-groupe maximal de G, puisque la maximalité est préservée par automorphisme. Cette réunion est alors disjointe d'après le fait 5.

Enfin, notons que, puisque la conjugaison par $g \in G$ est un automorphisme, on a

$$\big|g_i U g_i^{-1} \setminus \{1_G\}\big| \;=\; |U \setminus \{1_G\}| \;=\; |U| - 1.$$

On a donc finalement

$$\Big|\Big(\bigcup_{i=1}^{r} g_i U g_i^{-1}\Big) \setminus \{1_G\}\Big| = [G : U](|U| - 1) = \frac{n}{|U|}(|U| - 1) = n - \frac{n}{|U|}.$$

On va enfin pouvoir conclure. Soit U_1 un sous-groupe maximal de G. Il existe alors au moins un élément $x \in G$ non trivial qui n'appartient à aucun

conjugué de U_1. En effet, d'après le fait 7, le nombre de tels éléments est égal à

$$n - 1 - \Big(n - \frac{n}{|U_1|}\Big) = \frac{n}{|U_1|} - 1.$$

Or, cet entier est strictement positif puisque U_1 n'est pas G tout entier. D'après le lemme XI-2.8, il existe un sous-groupe maximal U_2 de G contenant $\langle x\rangle$. En effet, $\langle x\rangle$ est un sous-groupe propre de G puisque G n'est pas cyclique. En particulier, U_2 contient x et ne peut donc être conjugué à U_1, par choix de x. On en déduit alors aisément que les conjugués de U_1 et les conjugués de U_2 sont des sous-groupes maximaux deux à deux distincts.

Bref, les ensembles $\Big(\bigcup_{g\in G} gU_1g^{-1}\Big)\setminus\{1_G\}$ et $\Big(\bigcup_{g\in G} gU_2g^{-1}\Big)\setminus\{1_G\}$ sont disjoints. Comme il y a $n-1$ éléments distincts de 1_G dans G, le fait 7 nous donne l'inégalité

$$n - \frac{n}{|U_1|} + n - \frac{n}{|U_2|} < n,$$

soit encore

$$|U_1|\times|U_2| < |U_1| + |U_2|.$$

On a donc $|U_1|\times|U_2| - |U_1| - |U_2| = \big(|U_1|-1\big)\times\big(|U_2|-1\big) - 1 < 0$, ce qui est absurde car U_1 et U_2 possèdent au moins deux éléments.

Cela achève la démonstration. □

Pour finir, remarquons que n vérifie les conditions du théorème si, et seulement si, n et $\varphi(n)$ sont premiers entre eux.

En effet, si n vérifie la condition (C), on a déjà vu lors de la démonstration du fait 6 que n et $\varphi(d)$ était premiers entre eux pour tout diviseur d de n. En particulier, n et $\varphi(n)$ sont premiers entre eux.

Si maintenant, n et $\varphi(n)$ sont premiers entre eux, et si p est un diviseur premier de n, alors n ne peut être divisible par p^2, car sinon $\varphi(n)$ serait divisible par p. Par conséquent, on a $n = p_1\times\cdots\times p_r$, avec $p_1 < \cdots < p_r$, où chaque p_i est un nombre premier. S'il existait $i<j$ tels que $p_j \equiv 1\ [p_i]$, alors on aurait $p_i \mid p_j - 1$, et ainsi $p_j \mid \varphi(n)$. On en conclut que n doit vérifier la condition (C). Le théorème principal peut alors se reformuler comme suit.

3.2. Théorème. *Soit $n \geqslant 1$ un entier. Alors, tout groupe d'ordre n est cyclique si, et seulement si, n et $\varphi(n)$ sont premiers entre eux.*

Chapitre XIV

Petite classification des groupes finis ; épisodes précédents

Bien que nous ne soyons pas allés très loin dans le problème de classification des groupes finis, il nous a paru intéressant néanmoins de résumer les quelques résultats obtenus dans ce sens, aussi bien dans le cours que dans les exercices, ainsi que la classification des groupes finis d'ordre inférieur ou égal à 15.

Ce court chapitre, à vocation purement informative, ne contient donc pas d'exercices.

1. Quelques théorèmes de classification

On commence par rappeler le théorème de structure des groupe abéliens finis (théorème XII-5.3).

1.1. Théorème. *Soit G un groupe abélien fini.*

Alors, il existe des entiers $d_1, \ldots, d_s \geqslant 2$ vérifiant $d_1 \mid d_2 \mid \cdots \mid d_s$ tels que

$$G \simeq \mathbb{Z}/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_s\mathbb{Z}.$$

De plus, la suite d'entiers $(d_1, \ldots, d_s)$ est unique, et ne dépend que de la classe d'isomorphisme de G.

On continue par la structure des groupes d'ordre p et p^2, où p est un nombre premier (cf. corollaire VI-6.4 et exercice VII-10).

1.2. Théorème. *Soit p un nombre premier. Alors :*

(1) *tout groupe d'ordre p est cyclique, isomorphe à $\mathbb{Z}/p\mathbb{Z}$;*

(2) *tout groupe d'ordre p^2 est abélien. En particulier, un tel groupe est isomorphe à $\mathbb{Z}/p^2\mathbb{Z}$ ou à $(\mathbb{Z}/p\mathbb{Z})^2$.*

On donne maintenant la classification des groupes d'ordre pq, où p et q sont deux nombres premiers distincts (cf. exercice XI-4).

1.3. Théorème. *Soient p, q deux nombres premiers, avec $p < q$. Alors :*

(1) *si $p \nmid q-1$, tout groupe d'ordre pq est cyclique, isomorphe à $\mathbb{Z}/pq\mathbb{Z}$;*

(2) *si $p \mid q-1$, tout groupe d'ordre pq est soit isomorphe à $\mathbb{Z}/pq\mathbb{Z}$, soit isomorphe au produit semi-direct $\mathbb{Z}/q\mathbb{Z} \times_\rho \mathbb{Z}/p\mathbb{Z}$, où le morphisme de groupes[1] $\rho : \mathbb{Z}/p\mathbb{Z} \longrightarrow \mathrm{Aut}(\mathbb{Z}/q\mathbb{Z})$ est uniquement déterminé par*

$$\rho_{\bar{1}}(\tilde{x}) = \tilde{k}\tilde{x}, \qquad \text{pour tout } \tilde{x} \in \mathbb{Z}/q\mathbb{Z},$$

où $k \in \mathbb{Z}$ vérifie $q = 1 + kp$.

En prenant $p = 2$ et $q = \ell$, on en déduit le corollaire suivant (cf. également exercice X-9).

1.4. Corollaire. *Soit ℓ un nombre premier* **impair**. *Alors, tout groupe d'ordre 2ℓ est isomorphe à $\mathbb{Z}/2\ell\mathbb{Z}$ ou au groupe diédral D_ℓ.*

On a également le résultat suivant (cf. théorème XIII-3.2).

1.5. Théorème. *Soit $n \geqslant 1$ un entier. Alors, tout groupe d'ordre n est cyclique si, et seulement si, n et $\varphi(n)$ sont premiers entre eux.*

Nous avons bien entendu obtenu dans les exercices d'autres résultats de classification pour des groupes d'ordre spécifique. Nous en rappellerons certains dans le paragraphe suivant.

Achevons cette section en citant les théorèmes de classification sur les groupes finis simples obtenus dans cette partie. On commence par les groupes alternés.

1. En fait, tous les produits semi-directs correspondant à un morphisme non trivial sont isomorphes.

1.6. Théorème. *Soit E un ensemble fini non vide, possédant au moins trois éléments. Alors, le groupe alterné $\mathfrak{A}(E)$ est simple si, et seulement si, on a $|E| \neq 4$.*

On a aussi obtenu tous les groupes simples d'ordre 60 (cf. exercice XI-8).

1.7. Théorème. *Tout groupe simple d'ordre 60 est isomorphe à $\mathfrak{A}_5$.*

1.8. Remarque. Le groupe $\mathfrak{A}_5$ est en fait le plus petit groupe fini simple non abélien (cf. exercice XI-9).

Le théorème XI-2.22 donne immédiatement le résultat suivant.

1.9. Théorème. *Soit K un corps fini, et soit E un K-espace vectoriel de dimension finie $n \geqslant 2$. On suppose que $|K| \geqslant 4$ si $n = 2$. Alors, le groupe projectif spécial linéaire $\mathrm{PSL}(E)$ est simple.*

1.10. Remarque. Comme $\mathbb{F}_2^\times = \{\bar{1}\}$, on a, au vu des définitions,

$$\mathrm{GL}_3(\mathbb{F}_2) = \mathrm{SL}_3(\mathbb{F}_2) \simeq \mathrm{PSL}(\mathbb{F}_2^3).$$

En particulier, $\mathrm{GL}_3(\mathbb{F}_2)$ est simple[2] d'ordre 168.
On peut démontrer que c'est l'unique groupe simple d'ordre 168 à isomorphisme près (cf. [36], Exercice IV.21).

2. Groupes finis d'ordre $\leqslant 15$

Donnons maintenant la liste des groupes d'ordre $n \leqslant 15$ à isomorphisme près. Les cas $n \neq 8, 12$ se traitent en utilisant les théorèmes exposés dans le paragraphe précédent. Les deux cas restants sont respectivement résolus dans les exercices VI-45 et XI-17.

Dans la liste suivante, Q_8 est le groupe défini dans l'exercice VI-11, D_n est le groupe diédral d'ordre $2n$ (cf. exercice VI-40), et $\rho : \mathbb{Z}/4\mathbb{Z} \longrightarrow \mathrm{Aut}(\mathbb{Z}/3\mathbb{Z})$ est l'unique morphisme tel que

$$\rho_{\bar{1}}(\tilde{x}) = -\tilde{x}, \quad \text{pour tout } \tilde{x} \in \mathbb{Z}/4\mathbb{Z}.$$

- $n = 1$: $\{1\}$;
- $n = 2$: $\mathbb{Z}/2\mathbb{Z}$;
- $n = 3$: $\mathbb{Z}/3\mathbb{Z}$;
- $n = 4$: $\mathbb{Z}/4\mathbb{Z}$, $(\mathbb{Z}/2\mathbb{Z})^2$;

2. On peut aussi démontrer ce fait directement.

▷ $n = 5$: $\mathbb{Z}/5\mathbb{Z}$;

▷ $n = 6$: $\mathbb{Z}/6\mathbb{Z}$, D_3 ;

▷ $n = 7$: $\mathbb{Z}/7\mathbb{Z}$;

▷ $n = 8$: $\mathbb{Z}/8\mathbb{Z}$, $\mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, $(\mathbb{Z}/2\mathbb{Z})^3$, D_4, Q_8 ;

▷ $n = 9$: $\mathbb{Z}/9\mathbb{Z}$, $(\mathbb{Z}/3\mathbb{Z})^2$;

▷ $n = 10$: $\mathbb{Z}/10\mathbb{Z}$, D_5 ;

▷ $n = 11$: $\mathbb{Z}/11\mathbb{Z}$;

▷ $n = 12$: $\mathbb{Z}/12\mathbb{Z}$, $\mathbb{Z}/6\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$, D_6, $\mathfrak{A}_4$, $\mathbb{Z}/3\mathbb{Z} \times_\rho \mathbb{Z}/4\mathbb{Z}$;

▷ $n = 13$: $\mathbb{Z}/13\mathbb{Z}$;

▷ $n = 14$: $\mathbb{Z}/14\mathbb{Z}$, D_7 ;

▷ $n = 15$: $\mathbb{Z}/15\mathbb{Z}$.

2.1. Remarques

(1) Tout étudiant devrait connaître par cœur (voire savoir établir) la liste des groupes d'ordre $\leqslant 8$.

(2) La liste des groupes d'ordre p^3, où p est un nombre premier impair, est connue (cf. par exemple [24, Proposition XI.4.2]).

(3) Une liste des groupes d'ordre $\leqslant 32$ est proposée dans [24] (sans démonstration).

Les cas des groupes d'ordre 17, 19, 21, 22, 23, 25, 26, 29 et 31 peuvent être réglés grâce aux résultats évoqués dans le premier paragraphe, et une liste des groupes d'ordre 18 est établie dans l'exercice XI-19.

Le cas des groupes d'ordre 16 est déjà pénible : il y a cinq groupes abéliens et neuf groupes non abéliens. Pour les groupes d'ordre 32, c'est encore pire, puisqu'il y a sept groupes abéliens et quarante quatre groupes non abéliens.

Troisième partie

Anneaux : la loi des irréductibles

Chapitre XV

Propriétés élémentaires des anneaux et des corps

1. Généralités

Rappelons la définition d'anneau, déjà introduite dans le chapitre I (cf. définition I-3.15).

1.1. Définition. Un *anneau* est un triplet $(A, +, \cdot)$, où A est un ensemble non vide, $+ : A \times A \longrightarrow A$ et $\cdot : A \times A \longrightarrow A$ sont deux lois internes, vérifiant les propriétés suivantes :

(1) le couple $(A, +)$ est un groupe abélien ;

(2) la loi « $\cdot$ » est associative, et possède un élément neutre ;

(3) la loi « $\cdot$ » est distributive par rapport à « $+$ ».

Les éléments neutres pour les lois « $+$ » et « $\cdot$ » seront notés 0_A et 1_A respectivement.

On dit que l'anneau $(A, +, \cdot)$ est *commutatif* si « $\cdot$ » est commutative.

Attention ! $(A, \cdot)$ n'est pas un groupe multiplicatif (penser à $A = \mathbb{Z}$) !

1.2. Remarque. Le neutre 1_A pour la multiplication est unique d'après le lemme I-3.5.

1.3. Exemples

(1) Les ensembles $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ munis de leurs lois usuelles sont des anneaux commutatifs.

(2) Pour tout $n \geqslant 0$, l'ensemble $\mathbb{Z}/n\mathbb{Z}$, muni des lois

$$\begin{array}{ccc} \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \longrightarrow \mathbb{Z}/n\mathbb{Z} & & \mathbb{Z}/n\mathbb{Z} \times \mathbb{Z}/n\mathbb{Z} \longrightarrow \mathbb{Z}/n\mathbb{Z} \\ (\overline{a}, \overline{b}) \longmapsto \overline{a+b} & \text{et} & (\overline{a}, \overline{b}) \longmapsto \overline{ab} \end{array}$$

est un anneau commutatif, d'après la proposition II-6.3.

(3) Soit A un anneau et soit $n \geqslant 1$ un entier. Soit $\mathrm{M}_n(A)$ l'ensemble des matrices $n \times n$ à coefficients dans A. Si $M = (a_{ij})$, et $N = (b_{ij})$, on pose

$$M + N = (a_{ij} + b_{ij}) \quad \text{et} \quad MN = \Big(\sum_{k=1}^{n} a_{ik} b_{kj}\Big).$$

On obtient un anneau non commutatif si $n \geqslant 2$ (même si A l'est).

(4) Si $A_1, \ldots, A_n$ sont des anneaux, alors $A_1 \times \cdots \times A_n$ est un anneau pour les lois

$$(a_1, \ldots, a_n) + (a'_1, \ldots, a'_n) = (a_1 + a'_1, \ldots, a_n + a'_n)$$
$$(a_1, \ldots, a_n)(a'_1, \ldots, a'_n) = (a_1 a'_1, \ldots, a_n a'_n),$$

appelé *produit direct* des anneaux $A_1, \ldots, A_n$.
Les éléments neutres pour l'addition et la multiplication sont respectivement $(0_{A_1}, \ldots, 0_{A_n})$ et $(1_{A_1}, \ldots, 1_{A_n})$.

(5) Si X est un ensemble et A est un anneau, alors l'ensemble A^X des fonctions $f : X \to A$ est un anneau pour les lois suivantes : pour toutes fonctions $f, g \in A^X$, on définit les fonctions $f + g$ et fg par

$$(f+g)(x) = f(x) + g(x), \quad \text{et} \quad (fg)(x) = f(x)g(x), \quad \text{pour tout } x \in X.$$

(6) Soit $d \in \mathbb{Z} \setminus \{0\}$ un entier sans facteurs carrés. On définit $\sqrt{d}$ comme étant l'unique racine carrée positive de d dans $\mathbb{R}$ si $d > 0$, et comme égale à $i\sqrt{-d}$ si $d < 0$.
Soient $\mathbb{Z}[\sqrt{d}]$ et $\mathbb{Q}[\sqrt{d}]$ les sous-ensembles de $\mathbb{C}$ définis par :

$$\mathbb{Z}[\sqrt{d}] = \{a + b\sqrt{d} \in \mathbb{C} \mid a, b \in \mathbb{Z}\},$$
$$\mathbb{Q}[\sqrt{d}] = \{a + b\sqrt{d} \in \mathbb{C} \mid a, b \in \mathbb{Q}\}.$$

Alors, $\mathbb{Z}[\sqrt{d}]$ et $\mathbb{Q}[\sqrt{d}]$ sont des anneaux commutatifs pour les lois de l'anneau $\mathbb{C}$.
De plus, si $d \equiv 1\ [4]$, il en est de même de

$$\mathbb{Z}\Big[\frac{-1+\sqrt{d}}{2}\Big] = \Big\{a + b\frac{-1+\sqrt{d}}{2} \;\Big|\; a, b \in \mathbb{Z}\Big\}.$$

La vérification de ces faits est laissée en exercice.

Attention ! La dernière partie est fausse si $d \not\equiv 1[4]$. On laisse le soin au lecteur de trouver un contre-exemple.

Par exemple, les ensembles

$$\mathbb{Z}[i],\ \mathbb{Z}[\sqrt{2}],\ \mathbb{Z}[i\sqrt{2}],\ \mathbb{Z}[j],\ \mathbb{Q}[i],\ \mathbb{Q}[\sqrt{2}],\ \mathbb{Q}[i\sqrt{2}],\ \mathbb{Q}[j]$$

sont des anneaux commutatifs pour les lois de $\mathbb{C}$ $\left(\text{où } j = -\frac{1}{2} + i\frac{\sqrt{3}}{2}\right)$.

Le résultat suivant montre que les résultats familiers comme les règles de signes sont valables dans n'importe quel anneau.

1.4. Proposition. *Soit A un anneau. Pour tous $a, b \in A$, on a*

(1) $a\ 0_A = 0_A = 0_A\ a$;

(2) $a(-b) = -(ab) = (-a)b$;

(3) $(-1_A)a = -a$;

(4) $-(-a) = a$;

(5) $(-a)(-b) = ab$;

(6) $(-1_A)(-1_A) = 1_A$.

De plus, si A est commutatif, pour tous $a, b \in A$, et pour tout entier $n \geqslant 0$, on a

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}.$$

Démonstration. Par exemple, pour (1), puisque $0_A = 0_A + 0_A$, on

$$a\ 0_A = a(0_A + 0_A) = a\ 0_A + a\ 0_A.$$

On a donc $0_A = a\ 0_A + \big(-(a\ 0_A)\big) = a\ 0_A + a\ 0_A + \big(-(a\ 0_A)\big) = a\ 0_A$. La démonstration des autres identités est laissée au lecteur. □

Cette proposition implique le fait suivant : soit A un anneau pour lequel on a $0_A = 1_A$. Dans ce cas, A possède exactement un élément, puisque pour tout $a \in A$, on a

$$a = a\ 1_A = a\ 0_A = 0_A.$$

1.5. Définition. Un anneau A dans lequel $1_A = 0_A$ est appelé *anneau trivial.*

Comme pour le cas des groupes, il peut être pénible de montrer qu'un ensemble est un anneau. Heureusement, on a la notion de sous-anneau, qui joue le même rôle que celui de sous-groupe.

1.6. Définition. Soit A un anneau. Un *sous-anneau* de A est un sous-ensemble B de A vérifiant les conditions suivantes :

(1) on a $1_A \in B$;

(2) pour tous $x, y \in B$, $x+y, xy, -x \in B$.

Dans ce cas, l'axiome (2) montre que les lois internes de A induisent par restriction deux applications $B \times B \longrightarrow B$. De plus, on a $0_A \in B$. En effet, on a $1_A \in B$ par l'axiome (1), et donc $0_A = 1_A - 1_A \in B$ par l'axiome (2). Il est alors immédiat que B muni de ces deux opérations est un anneau, de neutres $0_B = 0_A$ et $1_B = 1_A$.

Autrement dit, un sous-anneau de A est un anneau pour les lois induites par celles de A.

Comme pour le cas des groupes, pour montrer qu'un ensemble muni de certaines opérations est un anneau, il est plus rapide de montrer que c'est un sous-anneau d'un anneau connu.

1.7. Exemples

(1) L'ensemble $\mathbb{Z}[\sqrt{d}]$ est un sous-anneau de $\mathbb{C}$.

De même, si $d \equiv 1$ [4], alors $\mathbb{Z}\Big[\dfrac{-1+\sqrt{d}}{2}\Big]$ est un sous-anneau de $\mathbb{C}$.

(2) L'anneau $\mathbb{Z}$ est un sous-anneau de $\mathbb{Q}$.

(3) L'ensemble $\mathrm{M}_n(\mathbb{Z})$ est un sous-anneau de $\mathrm{M}_n(\mathbb{Q})$.

(4) Plus généralement, si B est un sous-anneau de A, alors $\mathrm{M}_n(B)$ est un sous-anneau de $\mathrm{M}_n(A)$.

(5) L'ensemble $C^0([0,1],\mathbb{R})$ des fonctions continues $f : [0,1] \longrightarrow \mathbb{R}$ est un sous-anneau de l'anneau des fonctions de $[0,1]$ dans $\mathbb{R}$.

(6) L'ensemble $Z(A) = \{z \in A \mid az = za$, pour tout $a \in A\}$ est un sous-anneau de A, appelé le *centre* de A. [1]

La démonstration du résultat suivant est laissée en exercice.

1.8. Proposition. *Soit A un anneau. L'intersection d'une famille quelconque de sous-anneaux de A est encore un sous-anneau de A.*

Maintenant que nous avons défini les objets avec lesquels nous allons travailler, il reste à définir les morphismes entre les objets.

1. Attention ! C'est bien le produit de l'anneau qui est utilisé ici. Ainsi, $Z(A)$ n'est pas défini comme le centre du groupe $(A,+)$, ce qui n'aurait d'ailleurs aucun intérêt puisque, $(A,+)$ étant abélien, on aurait alors $Z(A) = A$.

1.9. Définition. Un *morphisme d'anneaux* $f : A \longrightarrow B$ est une application satisfaisant :

(1) $f(1_A) = 1_B$;

(2) pour tous $a, a' \in A$, $f(a + a') = f(a) + f(a')$;

(3) pour tous $a, a' \in A$, $f(aa') = f(a)f(a')$.

On note $\mathrm{Hom}(A, B)$ l'ensemble des morphismes d'anneaux $f : A \longrightarrow B$. Puisque $(B, +)$ est abélien, $\mathrm{Hom}(A, B)$ est un groupe abélien. On le notera aussi parfois $\mathrm{Hom}_{\mathrm{Ann}}(A, B)$ s'il y a risque de confusion avec l'ensemble des morphismes de groupes de A vers B. Clairement, la composition de deux morphismes d'anneaux est encore un morphisme d'anneaux.

Un *isomorphisme d'anneaux* $f : A \longrightarrow B$ est un morphisme d'anneaux bijectif. Dans ce cas, on vérifie comme dans la remarque VI-1.10 (2) que l'application inverse est aussi un isomorphisme d'anneaux. On dit que deux anneaux A et B sont isomorphes s'il existe un isomorphisme de A sur B. On note cela $A \simeq B$.

1.10. Remarques

(1) Un morphisme d'anneaux $f : A \longrightarrow B$ est un morphisme de groupes additifs. En particulier, on a toujours $f(0_A) = 0_B$.

(2) Si $f : A \longrightarrow B$ est un morphisme d'anneaux, alors, pour tout sous-anneau A' de A, $f(A')$ est un sous-anneau de B. En particulier, $\mathrm{Im}(f)$ est un sous-anneau de B.

En effet, $f(A')$ est un sous-groupe de B par le lemme VI-2.5 (1). On a donc $0_B \in f(A')$, et $f(A')$ est stable par addition et par opposé. De plus, puisque A' est un sous-anneau de A, on a $1_A \in A'$, et par conséquent, $1_B = f(1_A) \in f(A')$. Enfin, pour tous a'_1, $a'_2 \in A'$, on a

$$f(a'_1)f(a'_2) = f(a'_1a'_2) \in f(A'),$$

car $a'_1a'_2 \in A'$. Ainsi, $f(A')$ est stable par multiplication et $f(A')$ est un sous-anneau de B.

Les morphismes d'anneaux de $\mathbb{Z}$ vers un anneau A sont bien connus, et il n'y en a pas beaucoup !

1.11. Lemme. *Soit A un anneau. Alors, il existe un unique morphisme d'anneaux $\Theta_A : \mathbb{Z} \longrightarrow A$. Il est défini par*

$$\Theta_A(m) = m{\cdot}1_A, \qquad \textit{pour tout } m \in \mathbb{Z}.$$

Démonstration. Si $f : \mathbb{Z} \longrightarrow A$ est un morphisme d'anneaux, on a néces-

sairement $f(1) = 1_A$, et par suite

$$f(m) = f(1 + \cdots + 1) = f(1) + \cdots + f(1) = m{\cdot}1_A, \qquad \text{pour tout } m \geqslant 0.$$

Si $m \leqslant 0$, on a alors

$$f(m) = f\big(-(-m)\big) = -f(-m) = -(-m{\cdot}1_A) = m{\cdot}1_A,$$

par la proposition 1.4. Ainsi, on a nécessairement $f = \Theta_A$. Il reste à vérifier que Θ_A est bien un morphisme d'anneaux, ce qui est un exercice longuet, mais facile, lâchement laissé au lecteur. □

Le noyau de cet unique morphisme va maintenant nous permettre de définir un entier très important attaché à un anneau A, à savoir, sa caractéristique.

1.12. Définition. Soit A un anneau. Le morphisme d'anneaux

$$\begin{aligned} \Theta_A : \mathbb{Z} &\longrightarrow A \\ m &\longmapsto m{\cdot}1_A. \end{aligned}$$

est en particulier un morphisme de groupes abéliens. Son noyau est ainsi un sous-groupe de $\mathbb{Z}$, donc de la forme $c\mathbb{Z}$ d'après la proposition VI-2.4.

La *caractéristique* de A est l'unique entier $c \geqslant 0$ défini par

$$\mathrm{Ker}(\Theta_A) = c\mathbb{Z}.$$

On la note $\mathrm{car}(A)$. Autrement dit, $\mathrm{car}(A) > 0$ si, et seulement si, l'équation

$$m{\cdot}1_A = 0_A, \quad m \in \mathbb{Z}$$

possède au moins une solution non nulle. Dans ce cas, $\mathrm{car}(A)$ est le plus petit entier $c > 0$ tel que $c{\cdot}1_A = 0_A$. De plus, pour tout $m \in \mathbb{Z}$, on a

$$m{\cdot}1_A = 0_A \quad \Longrightarrow \quad \mathrm{car}(A) \mid m.$$

1.13. Remarque. La caractéristique de A est non nulle si, et seulement si, 1_A est d'ordre fini dans $(A, +)$. Dans ce cas, on a $\mathrm{car}(A) = o(1_A)$.

Cela provient simplement des définitions.

1.14. Exemple. Les anneaux $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$, $\mathbb{C}$ sont de caractéristique 0, tandis que $\mathrm{car}(\mathbb{Z}/n\mathbb{Z}) = n$ pour tout $n \geqslant 1$.

1.15. Remarque. Si B est un sous-anneau de A, alors A et B ont même caractéristique.

Cela provient de la définition, et du fait que $1_B = 1_A$.

On a alors la propriété importante suivante.

1.16. Proposition. *Soit A un anneau. Alors, A contient un sous-anneau isomorphe à $\mathbb{Z}/\mathrm{car}(A)\mathbb{Z}$.*

Démonstration. Le théorème de factorisation pour les groupes appliqué au morphisme Θ_A montre que l'on a un morphisme de groupes injectif

$$\overline{\Theta}_A : \mathbb{Z}/\mathrm{car}(A)\mathbb{Z} \hookrightarrow A.$$

On vérifie aisément que c'est aussi un morphisme d'anneaux. □

1.17. Remarque. Si $c = \mathrm{car}(A)$, alors, pour tout $m \in \mathbb{Z}$, on a

$$c \mid m \quad \Longrightarrow \quad m{\cdot}a = 0_A, \ \text{ pour tout } a \in A.$$

En effet, soit $m \in \mathbb{Z}$ tel que $c \mid m$.

Supposons tout d'abord que $m \geqslant 0$. Si $a \in A$, on a

$$m{\cdot}a = a + \cdots + a = a(1_A + \cdots + 1_A) = a(m{\cdot}1_A) = a0_A = 0_A.$$

Si $m \leqslant 0$, on a alors

$$m{\cdot}a = (-a) + \cdots + (-a) = (-m){\cdot}(-a) = 0_A,$$

d'où l'assertion.

On en déduit aisément que, si $m{\cdot}1_A = 0_A$ a une solution non nulle dans $\mathbb{Z}$, alors $\mathrm{car}(A)$ est le plus petit entier $c > 0$ tel que $c{\cdot}a = 0_A$ pour tout $a \in A$.

Nous terminons ce paragraphe par un lemme calculatoire très utile.

1.18. Lemme. *Soit A un anneau commutatif de caractéristique p, où p est un nombre premier. Alors, pour tout $m \geqslant 0$, on a*

$$(a+b)^{p^m} = a^{p^m} + b^{p^m}, \quad (ab)^{p^m} = a^{p^m} b^{p^m}, \quad \textit{pour tous } a, b \in A.$$

En particulier, l'application

$$\begin{aligned} A &\longrightarrow A \\ a &\longmapsto a^{p^m} \end{aligned}$$

est un morphisme d'anneaux.

Démonstration. Le seul point non évident est la première égalité. Il suffit de traiter le cas $m = 1$, le cas général s'en déduisant par une récurrence évidente. Puisque A est commutatif, pour tous $a, b \in A$, on a

$$(a+b)^p = \sum_{k=0}^{p} \binom{p}{k} a^k b^{p-k}.$$

Or, il est bien connu que, pour tout $k \in [\![1, p-1]\!]$, $p \mid \binom{p}{k}$. On applique alors la remarque précédente pour conclure. □

2. Éléments remarquables d'un anneau

L'existence d'une loi produit dans un anneau permet de définir quelques familles d'éléments remarquables. On commence par les éléments inversibles.

2.1. Définition. Soit A un anneau. On dit que $a \in A$ est *inversible* (ou une *unité*) s'il possède un symétrique pour la loi produit, i.e. s'il existe $b \in A$ tel que $ab = 1_A = ba$.

Dans ce cas, ce symétrique est unique d'après le lemme I-3.8. Il est noté a^{-1} et est appelé *inverse* de a.

L'ensemble des éléments inversibles de A est noté $A^\times$.

2.2. Proposition. *Soit A un anneau. Alors, $A^\times$ est un groupe pour la multiplication. De plus, pour tous $a, b \in A^\times$, on a*

$$(a^{-1})^{-1} = a, \qquad (ab)^{-1} = b^{-1}a^{-1}.$$

Démonstration. Remarquons que $1_A \in A^\times$, puisque $1_A\, 1_A = 1_A$. Ainsi, $A^\times$ est non vide. Par ailleurs, si $a, b \in A^\times$, alors

$$(ab)b^{-1}a^{-1} = a(bb^{-1})a^{-1} = a\ 1_A a^{-1} = aa^{-1} = 1_A.$$

De même, $b^{-1}a^{-1}(ab) = 1_A$. Par conséquent, ab est bien inversible, d'inverse $b^{-1}a^{-1}$, et la loi de multiplication sur A se restreint donc en une loi interne sur $A^\times$. Puisque A est un anneau, cette loi est associative, de neutre 1_A. Enfin, si $a \in A^\times$, alors $a^{-1} \in A$ est aussi inversible, d'inverse a, puisque

$$a^{-1}a = aa^{-1} = 1_A.$$

Ainsi, si $a \in A^\times$, on a $a^{-1} \in A^\times$ et $(a^{-1})^{-1} = a$. De plus, a^{-1} est clairement un inverse de a dans $A^\times$. Cela achève la démonstration. □

2.3. Exemples

(1) Dans un anneau A non trivial, 0_A n'est jamais inversible. En particulier, si A est non trivial, on a $A^\times \subset A \setminus \{0_A\}$.

Par ailleurs, 1_A et -1_A sont toujours inversibles dans n'importe quel anneau A, donc $\{\pm 1_A\} \subset A^\times$. Attention néanmoins, ces deux éléments ne sont pas nécessairement distincts (penser par exemple à $A = \mathbb{Z}/2\mathbb{Z}$).

(2) On a $2 \in \mathbb{Q}^\times$, mais $2 \notin \mathbb{Z}^\times$. En effet, puisque $\mathbb{Z} \subset \mathbb{Q}$, si 2 a un inverse dans $\mathbb{Z}$, il a un inverse dans $\mathbb{Q}$, qui est nécessairement $\frac{1}{2}$. En revanche, on a $\frac{1}{2} \notin \mathbb{Z}$. En fait, on a $\mathbb{Z}^\times = \{\pm 1\}$.

(3) On a $\mathbb{C}^\times = \mathbb{C} \setminus \{0\}$.

2.4. Remarque. Tout morphisme d'anneaux $f : A \longrightarrow B$ induit par restriction un morphisme de groupes $f : A^\times \longrightarrow B^\times$.

En effet, si $a \in A^\times$, on a

$$f(a)f(a^{-1}) = f(aa^{-1}) = f(1_A) = 1_B,$$

et de même, on a

$$f(a^{-1})f(a) = f(a^{-1}a) = f(1_A) = 1_B.$$

Ainsi, $f(a) \in B^\times$ pour tout $a \in A^\times$. Par conséquent, f restreint en une application $f : A^\times \longrightarrow B^\times$. C'est un morphisme de groupes puisque un morphisme d'anneaux respecte les lois de multiplication.

La proposition 1.4 montre que les règles de calculs dans un anneau quelconque sont similaires aux règles de calculs bien connues sur les entiers ou les réels. Cela est vrai jusqu'à un certain point. Par exemple, dans $\mathbb{Z}$, le produit de deux facteurs est nul si, et seulement si, un des deux facteurs est nul.

Cela devient faux dans un anneau quelconque.

2.5. Exemples

(1) Dans $\mathbb{Z}/6\mathbb{Z}$, on a $\overline{2} \neq \overline{0}$ et $\overline{3} \neq \overline{0}$, mais $\overline{2}\cdot\overline{3} = \overline{6} = \overline{0}$.

(2) Dans $\mathrm{M}_2(\mathbb{R})$, on a $\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Cela motive la définition suivante.

2.6. Définition. Soit A un anneau. On dit que A possède des *diviseurs de zéro* s'il existe $a, b \in A \setminus \{0_A\}$ tels que $ab = 0_A$. Dans ce cas, a est appelé un *diviseur de zéro à gauche* et b est appelé un *diviseur de zéro à droite*.

Si A est commutatif, les notions de diviseur de zéro à gauche et à droite coïncident, et l'on parlera simplement de diviseur de zéro.

2.7. Exemples

(1) Si $A = \mathbb{C} \times \mathbb{C}$, alors $(1,0)$ et $(0,1)$ sont des diviseurs de zéro.

(2) Si $A = \mathbb{Z}/6\mathbb{Z}$, alors $\overline{2}$ et $\overline{3}$ sont des diviseurs de zéro.

(3) Si $A = \mathrm{M}_2(\mathbb{R})$, alors

$$M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

est un diviseur de zéro à gauche et à droite, puisque $M^2 = 0_{\mathrm{M}_2(\mathbb{R})}$.

Nous allons maintenant définir deux familles de diviseurs de zéro particuliers, qui seront étudiées plus en détail dans la suite, à savoir les éléments nilpotents, et les idempotents.

2.8. Définition. Soit A un anneau. On dit que $a \in A$ est *nilpotent* s'il existe $n \geqslant 1$ tel que $a^n = 0_A$.

L'ensemble des éléments nilpotents est noté $\mathrm{Nil}(A)$.

2.9. Remarque. Un élément nilpotent non nul est un diviseur de zéro à gauche et à droite.

En effet, soit $a \in \mathrm{Nil}(A) \setminus \{0_A\}$. Alors, il existe un entier $n \geqslant 1$ minimal tel que $a^n = 0_A$. Mais alors, $a^{n-1} \neq 0_A$ et $aa^{n-1} = a^{n-1}a = 0_A$.

2.10. Exemples

(1) Si $A = \mathbb{C} \times \mathbb{C}$, alors $\mathrm{Nil}(A) = \{(0,0)\}$.

(2) Si $A = \mathbb{Z}/4\mathbb{Z}$, alors $\mathrm{Nil}(A) = \{\overline{0}, \overline{2}\}$.

(3) Si $A = \mathrm{M}_2(\mathbb{R})$, alors

$$M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$$

est un élément nilpotent.

L'existence d'éléments nilpotents non nuls permet de construire des éléments inversibles différents de 1_A.

2.11. Lemme. *Soit A un anneau. On a*

$$\forall a \in \mathrm{Nil}(A), \quad -a \in \mathrm{Nil}(A) \ \text{ et } \ 1_A \pm a \in A^\times.$$

Démonstration. Soit $a \in \mathrm{Nil}(A)$, et soit $n \geqslant 1$ tel que $a^n = 0_A$. Alors, on a

$$(-a)^n = (-1_A)^n a^n = 0_A,$$

et ainsi $-a \in \mathrm{Nil}(A)$. De plus, on a

$$(1_A + a + \cdots + a^{n-1})(1_A - a) = (1_A - a)(1_A + a + \cdots + a^{n-1}) = 1_A - a^n = 1_A,$$

ce qui démontre que $1_A - a \in A^\times$.

En remplaçant a par $-a$, qui est aussi nilpotent, on obtient $1_A + a \in A^\times$. □

On continue cette fois avec les éléments idempotents.

2.12. Définition. Soit A un anneau. Un *idempotent* est un élément $e \in A$ tel que $e^2 = e$.

Les éléments 0_A et 1_A sont des idempotents, dit triviaux.

2.13. Remarques. Soit A un anneau.

(1) Si $e \in A$ est un idempotent non trivial, alors e et $1_A - e$ sont des diviseurs de zéro à gauche et à droite, puisque $e(1_A - e) = (1_A - e)e = 0$.

(2) Si $e \in A$ est un idempotent, alors $1_A - e$ est aussi un idempotent, puisque
$$(1_A - e)^2 = 1_A - 2e + e^2 = 1_A - e.$$

2.14. Exemples

(1) Les seuls idempotents d'un anneau sans diviseurs de zéro sont triviaux.

(2) Si $A = \mathbb{Z}/6\mathbb{Z}$, $\overline{3}$ est un idempotent non trivial.

(3) Si K est un corps, les idempotents de $\mathrm{M}_n(K)$ sont exactement les matrices de projecteurs.

Comme nous le verrons plus loin, les éléments idempotents interviennent naturellement lorsque l'on cherche à écrire un anneau donné comme produit direct d'anneaux. Nous nous contenterons pour l'instant de démontrer le lemme suivant.

2.15. Lemme. *Soit A un anneau. Pour tout idempotent $e \in A$, l'ensemble eAe est un anneau pour les lois de A, de neutre additif 0_A, et de neutre multiplicatif e.*

Démonstration. Pour tous a, $a' \in A$, on a $eae + ea'e = e(a + a')e \in eAe$, ainsi que $(eae)(ea'e) = eae^2a'e = e(aea')e \in eAe$. Par conséquent, les lois de A se restreignent en deux lois internes sur eAe. Les propriétés d'associativité et de distributivité sont alors conservées. De plus, on vérifie aisément que l'ensemble eAe est un sous-groupe de $(A, +)$, donc a fortiori un groupe abélien, de neutre 0_A. Il reste donc à voir que e est un neutre pour la loi produit. Or, pour tout $a \in A$, on a
$$e(eae) = e^2ae = eae,$$
et de même
$$(eae)e = eae^2 = eae,$$
d'où le résultat voulu. □

2.16. Remarque. Si $e \in A$ est un idempotent, eAe est un anneau contenu dans A, et qui n'est pas un sous-anneau de A dès que $e \neq 1_A$.

Les anneaux ne possédant pas de diviseur de zéro sont évidemment parmi les plus agréables, puisque les choses redeviennent plus naturelles du point de vue calculatoire. Ceux-ci méritent une attention particulière et font l'objet du paragraphe suivant.

3. Anneaux intègres, corps

3.1. Définition. Un anneau A est dit *intègre* s'il est non trivial, commutatif et ne possède pas de diviseurs de zéro. Autrement dit, A est intègre s'il est non trivial, commutatif, et pour tous $a, b \in A$, on a

$$ab = 0_A \implies a = 0_A \text{ ou } b = 0_A.$$

3.2. Remarque. Un sous-anneau d'un anneau intègre est un anneau intègre. La démonstration de ce fait est laissée en exercice.

3.3. Exemples

(1) Les anneaux $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$ sont intègres.

(2) Dans l'anneau $\mathbb{Z}/4\mathbb{Z}$, on a $\overline{2} \neq \overline{0}$, puisque 4 ne divise pas 2, mais on a par ailleurs $\overline{2}\,\overline{2} = \overline{4} = \overline{0}$, donc $\overline{2}$ est un diviseur de zéro, et $\mathbb{Z}/4\mathbb{Z}$ n'est pas un anneau intègre.

(3) Si A est non trivial, l'anneau $A \times A$ n'est pas intègre, puisque
$$(1_A, 0_A)\,(0_A, 1_A) = (0_A, 0_A).$$

(4) Les anneaux $\mathbb{Z}[\sqrt{d}]$ et $\mathbb{Z}\left[\frac{-1+\sqrt{d}}{2}\right]$ (lorsque $d \equiv 1$ [4]) sont intègres.

Revenons maintenant aux éléments inversibles d'un anneau. Comme on l'a déjà constaté, tout élément non nul d'un anneau n'est pas nécessairement inversible. Lorsque c'est le cas, on peut donc « diviser » par n'importe quel élément non nul, d'où la définition suivante.

3.4. Définition. Un anneau A est un *anneau à division* s'il est non trivial et si tout élément non nul est inversible.

Un *corps* est un anneau à division qui est commutatif. Autrement dit, un corps est un anneau non trivial, commutatif, tel que tout élément non nul est inversible.

3.5. Remarque. Le lecteur vérifiera que cette définition est en accord avec celle donné dans le chapitre I (cf. définition I-3.17).

3.6. Exemples

(1) L'anneau $\mathbb{Z}$ n'est pas un corps.

(2) Les anneaux $\mathbb{Q}$, $\mathbb{R}$ et $\mathbb{C}$ sont des corps.

(3) Les anneaux $\mathbb{Z}[\sqrt{d}]$ et $\mathbb{Z}\left[\frac{-1+\sqrt{d}}{2}\right]$ ne sont pas des corps.

(4) Pour tout $d \in \mathbb{Z} \setminus \{0\}$ sans facteurs carrés, l'anneau $\mathbb{Q}[\sqrt{d}]$ est un corps (exercice laissé au lecteur).

3.7. Remarque. L'exercice 20 donne un exemple d'un anneau à division qui n'est pas commutatif.

3.8. Proposition. *Un diviseur de zéro à gauche ou à droite n'est pas inversible. En particulier, un anneau à division ne contient pas de diviseurs de zéro, et tout corps est un anneau intègre.*

Démonstration. Soient $a, b \in A\backslash\{0_A\}$ tels que $ab = 0_A$. Si a était inversible, on aurait

$$b = 1_A\, b = (a^{-1}a)b = a^{-1}(ab) = a^{-1}0_A = 0_A,$$

d'où une contradiction, et de même pour b. Le reste est clair. □

Attention ! Un anneau intègre n'est pas nécessairement un corps (penser à l'anneau $\mathbb{Z}$).

Examinons maintenant le cas de l'anneau $\mathbb{Z}/n\mathbb{Z}$.

3.9. Proposition. *Soit $n \geqslant 1$ un entier. Alors, $\mathbb{Z}/n\mathbb{Z}$ est intègre si, et seulement si, n est premier. Dans ce cas, $\mathbb{Z}/n\mathbb{Z}$ est un corps.*

Démonstration. Si $n = 1$, alors n n'est pas premier. De plus, $\mathbb{Z}/1\mathbb{Z} = \{\overline{0}\}$. On a donc $\overline{0} = \overline{1}$, et donc $\mathbb{Z}/n\mathbb{Z}$ est trivial, et n'est pas un corps.

Supposons que $n \geqslant 2$. Si n est premier, la proposition II-6.7 montre que l'anneau $\mathbb{Z}/n\mathbb{Z}$ est un corps, et donc intègre par la proposition précédente.

Si n n'est pas premier, il existe $n_1, n_2 \in [\![1, n-1]\!]$ tels que $n = n_1 n_2$. Remarquons que les classes $\overline{n}_1$ et $\overline{n}_2$ sont non nulles par le lemme II-6.5, puisque $n_1, n_2 \in [\![1, n-1]\!]$. En revanche, on a

$$\overline{n}_1\overline{n}_2 = \overline{n_1 n_2} = \overline{n} = \overline{0}.$$

Ainsi, $\mathbb{Z}/n\mathbb{Z}$ n'est pas intègre, et cela achève la démonstration. □

3.10. Corollaire. *La caractéristique d'un anneau intègre est soit nulle, soit un nombre premier.*

Démonstration. Soit A un anneau intègre. Par la proposition 1.16, A contient un sous-anneau isomorphe à $\mathbb{Z}/\mathrm{car}(A)\mathbb{Z}$, qui est donc lui-même intègre. Or, $\mathbb{Z}/\mathrm{car}(A)\mathbb{Z}$ est intègre si, et seulement si, $\mathrm{car}(A) = 0$ (car cet anneau est isomorphe à $\mathbb{Z}$ dans ce cas) ou si $\mathrm{car}(A)$ est un nombre premier, par la proposition 3.9. Cela achève la démonstration. □

Attention ! La réciproque est fausse. En effet, $\mathbb{C} \times \mathbb{C}$ est un anneau de caractéristique 0, mais n'est pas intègre.

On a bien entendu la notion de sous-anneau à division et de sous-corps.

3.11. Définition. Soit A un anneau à division (resp. un corps). Un *sous-anneau à division* (resp. un *sous-corps*) est un sous-anneau B de A tel que, pour tout $x \in B \setminus \{0_A\}$, on a $x^{-1} \in B$.

Autrement dit, c'est un sous-ensemble contenant 0_A, 1_A, et stable par addition, multiplication, passage à l'opposé, et passage à l'inverse.

Dans ce cas, B est un anneau à division (resp. un corps) pour les lois induites par celles de A.

3.12. Exemple. L'anneau $\mathbb{Q}$ est un sous-corps de $\mathbb{R}$, qui est lui-même un sous-corps de $\mathbb{C}$.

Nous finissons ce paragraphe en généralisant la construction du corps des rationnels. La motivation de cette construction est la suivante. Si K est un corps, alors tout sous-anneau A de K est intègre. Inversement, si A est un anneau intègre, peut-on l'identifier à un sous-anneau d'un corps K_A bien choisi ? L'existence d'un tel corps K_A permettrait alors de faire des calculs plus facilement que dans A, puisque tout élément non nul de A aurait un inverse dans K_A, même s'il n'est pas inversible dans A, de la même manière qu'un entier relatif non nul est toujours inversible dans $\mathbb{Q}$.

Pour construire un tel corps K_A, l'idée est justement d'imiter la construction de $\mathbb{Q}$ à partir de l'anneau $\mathbb{Z}$.

Soit A un anneau intègre. On définit une relation sur $A \times (A \setminus \{0_A\})$ de la façon suivante. Si (a,b), $(a',b') \in A \times (A \setminus \{0_A\})$, on dit que $(a,b) \sim (a',b')$ si $ab' - ba' = 0_A$.

3.13. Lemme. *La relation $\sim$ est une relation d'équivalence.*

Démonstration. Soient (a,b), (a',b'), $(a'',b'') \in A \times (A \setminus \{0_A\})$.

Puisque A est commutatif, on a $ab - ba = 0_A$, et donc $(a,b) \sim (a,b)$. Ainsi, la relation $\sim$ est réflexive. Supposons maintenant que $(a,b) \sim (a',b')$. On a donc $ab' - ba' = 0_A$. Mais alors,

$$a'b - ab' = -(ab' - ba') = 0_A,$$

et par conséquent, $(a',b') \sim (a,b)$, d'où la symétrie de la relation. Enfin, supposons que $(a,b) \sim (a'b')$ et $(a',b') \sim (a'',b'')$, c'est-à-dire

$$ab' - ba' = a'b'' - b'a'' = 0_A.$$

On a alors

$$b'(ab'' - a''b) = b''ab' - bb'a'' = b''ba' - ba'b'' = 0_A.$$

Comme $b' \neq 0_A$ et A est un anneau intègre, on obtient $ab'' - a''b = 0_A$, c'est-à-dire $(a,b) \sim (a'',b'')$, d'où la transitivité. □

On note K_A l'ensemble quotient $A \times (A \setminus \{0_A\})/\sim$, et l'on note $\frac{a}{b}$ la classe d'équivalence de (a, b).

On définit aussi une loi d'addition et une loi de multiplication sur les éléments de $A \times (A \setminus \{0\})$ de la manière suivante :

$$(a, b) + (c, d) = (ad + bc, bd) \quad \text{et} \quad (a, b)\cdot(c, d) = (ac, bd).$$

Remarquons que la définition a bien un sens, puisque si b et d sont non nuls, alors bd est non nul par intégrité de A.

On a alors le lemme suivant.

3.14. Lemme. *Soient* (a, b), (a', b'), (c, d), $(c', d') \in (A \times (A \setminus \{0_A\}))$. *On suppose que* $(a, b) \sim (a', b')$ *et* $(c, d) \sim (c', d')$. *Alors, on a*

$$(a, b) + (c, d) \sim (a', b') + (c', d') \quad \text{et} \quad (a, b)\cdot(c, d) \sim (a', b')\cdot(c', d').$$

Démonstration. Laissée au lecteur. □

Ce lemme démontre que les lois sur $A \times (A \setminus \{0_A\})$ induisent des lois sur l'ensemble quotient K_A par passage au quotient.

Par définition, on a donc

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd} \quad \text{et} \quad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}.$$

3.15. Proposition. *L'ensemble* K_A, *muni des deux lois précédentes, est un corps commutatif, dont les éléments neutres pour les loi « + » et « · » sont* $0_{K_A} = \frac{0_A}{1_A}$ *et* $1_{K_A} = \frac{1_A}{1_A}$ *respectivement.*

De plus, on a

(1) *pour tout* $a \in A$, *et tous* $b, c \in A \setminus \{0_A\}$, *on a* $\frac{a}{b} = \frac{ac}{bc}$;

(2) *pour tout* $a \in A$, *et tout* $b \in A \setminus \{0_A\}$, *on a* $-\frac{a}{b} = \frac{-a}{b} = \frac{a}{-b}$;

(3) *pour tous* $a, b \in A \setminus \{0_A\}$, *on a* $\frac{a}{b} \neq 0_{K_A}$ *et son inverse multiplicatif est* $\frac{b}{a}$;

(4) *l'application*

$$\begin{aligned} \iota_A \colon A &\longrightarrow K_A \\ a &\longmapsto \frac{a}{1_A} \end{aligned}$$

est un morphisme d'anneaux injectif.

Démonstration. Exactement comme dans le cas du corps des rationnels. □

3.16. Définition. Soit A un anneau intègre. Le corps K_A s'appelle le *corps des fractions* de A.

3.17. Exemples

(1) Si $A = \mathbb{Z}$, on a $K_A = \mathbb{Q}$.

(2) Si A est un corps, alors ι_A induit un isomorphisme d'anneaux $K_A \simeq A$. Cela est laissé en exercice au lecteur.

La proposition suivante décrit la propriété universelle du corps des fractions.

3.18. Proposition. *Soit A un anneau intègre. Soit B un anneau commutatif, et soit $f : A \longrightarrow B$ un morphisme d'anneaux tel que $f(A\setminus\{0_A\}) \subset B^\times$. Alors, il existe un unique morphisme d'anneaux $f_A : K_A \longrightarrow B$ tel que*

$$f_A\Big(\frac{a}{1_A}\Big) = f(a), \quad \textit{pour tout } a \in A.$$

Ce morphisme est défini par

$$f_A\Big(\frac{a}{b}\Big) = f(a)f(b)^{-1}, \quad \textit{pour tout } \frac{a}{b} \in K_A.$$

Démonstration. Supposons tout d'abord qu'un tel morphisme f_A existe. Pour tout $\frac{a}{b} \in K_A$, on a

$$\frac{a}{b} = \frac{a}{1_A}\frac{1_A}{b} = \frac{a}{1_A}\Big(\frac{b}{1_A}\Big)^{-1},$$

et par conséquent

$$f_A\Big(\frac{a}{b}\Big) = f_A\Big(\frac{a}{1_A}\Big)f_A\Big(\frac{b}{1_A}\Big)^{-1} = f(a)f(b)^{-1}.$$

Ainsi, si un tel morphisme $f_A : K_A \longrightarrow B$ existe, il est unique, et défini par la formule ci-dessus. Il reste donc à vérifier que l'application

$$\begin{aligned} f_A\colon K_A &\longrightarrow B \\ \frac{a}{b} &\longmapsto f(a)f(b)^{-1} \end{aligned}$$

est bien définie et est un morphisme d'anneaux.

Soient (a,b), $(a',b') \in A\times(A\setminus\{0_A\})$ tels que $(a,b) \sim (a',b')$. Puisque l'on a l'égalité $ab' - ba' = 0_A$ et que f est un morphisme d'anneaux, on obtient

$$f(a)f(b') - f(b)f(a') = f(ab' - a'b) = f(0_A) = 0_B.$$

Par hypothèse, $f(b)$ et $f(b')$ sont inversibles. Comme B est commutatif, un calcul immédiat montre que l'égalité précédente se récrit

$$f(a)f(b)^{-1} = f(a')f(b')^{-1}.$$

Il s'ensuit que l'application

$$\begin{aligned} A \times A \setminus \{0_A\} &\longrightarrow B \\ (a,b) &\longmapsto f(a)f(b)^{-1} \end{aligned}$$

est constante sur les classes d'équivalence, et induit donc par passage au quotient une application

$$\begin{aligned} f_A : K_A &\longrightarrow B \\ \frac{a}{b} &\longmapsto f(a)f(b)^{-1}. \end{aligned}$$

Vérifions maintenant que f_A est un morphisme d'anneaux possédant la propriété voulue. Tout d'abord, pour tout $a \in A$, on a

$$f_A\Big(\frac{a}{1_A}\Big) = f(a)f(1_A)^{-1} = f(a)1_B^{-1} = f(a).$$

En particulier, $f_A(1_{K_A}) = 1_B$. De plus, pour tous $\frac{a}{b}, \frac{c}{d} \in K_A$, on a

$$\begin{aligned} f_A\Big(\frac{a}{b}\,\frac{c}{d}\Big) &= f_A\Big(\frac{ac}{bd}\Big) \\ &= f(ac)f(bd)^{-1} \\ &= f(a)f(c)f(d)^{-1}f(b)^{-1} \\ &= f(a)f(b)^{-1}f(c)f(d)^{-1} \\ &= f_A\Big(\frac{a}{b}\Big)f_A\Big(\frac{c}{d}\Big), \end{aligned}$$

ainsi que

$$\begin{aligned} f_A\Big(\frac{a}{b} + \frac{c}{d}\Big) &= f_A\Big(\frac{ad+bc}{bd}\Big) \\ &= f(ad+bc)f(bd)^{-1} \\ &= \big(f(a)f(d) + f(b)f(c)\big)f(d)^{-1}f(b)^{-1} \\ &= f(a)f(b)^{-1} + f(c)f(d)^{-1} \\ &= f_A\Big(\frac{a}{b}\Big) + f_A\Big(\frac{c}{d}\Big). \end{aligned}$$

Cela achève la démonstration. □

Nous revenons maintenant à des considérations plus générales.

4. Idéaux

Nous allons maintenant nous attarder un peu sur une notion cruciale en théorie des anneaux, et qui jouera le même rôle que les sous-groupes distingués d'un groupe en théorie des groupes. D'après le lemme VI-2.15, le noyau d'un morphisme de groupes $\varphi : G \longrightarrow G'$ est un sous-groupe distingué de G.

Par ailleurs, tout sous-groupe H distingué dans G est le noyau de la projection canonique $G \longrightarrow G/H$. Ainsi, les sous-groupes distingués de G sont exactement les sous-groupes apparaissant comme noyau d'un morphisme de groupes de source G.

Regardons maintenant les propriétés du noyau d'un morphisme d'anneaux. Outre le fait que le noyau d'un morphisme d'anneaux $f : A \longrightarrow B$ est un sous-groupe de $(A, +)$, il vérifie également la propriété suivante : pour tout $a \in A$, et pour tout $x \in \mathrm{Ker}(f)$, on a $xa \in \mathrm{Ker}(f)$ et $ax \in \mathrm{Ker}(f)$.

Cela motive la notion d'idéal (bilatère) d'un anneau. On verra plus loin que cette notion vient naturellement lorsque l'on cherche à construire les anneaux quotients, tout comme les sous-groupes distingués d'un groupe interviennent de façon naturelle dans la construction des groupes quotients.

Comme nous ne supposons pas l'anneau A commutatif, nous avons également une notion d'idéal à gauche (resp. d'idéal à droite), qui n'interviendra que de manière anecdotique dans ce chapitre. Néanmoins, par souci d'exhaustivité, nous en donnons la définition.

4.1. Définition. Soit A un anneau.

Un sous-ensemble $\mathfrak{a} \subset A$ est un *idéal à gauche* si

(1) $\mathfrak{a}$ est un sous-groupe additif de $(A, +)$;

(2) pour tout $a \in A$ et tout $x \in \mathfrak{a}$, on a $ax \in \mathfrak{a}$.

Un sous-ensemble $\mathfrak{a} \subset A$ est un *idéal à droite* si

(1) $\mathfrak{a}$ est un sous-groupe additif de $(A, +)$;

(2) pour tout $a \in A$ et tout $x \in \mathfrak{a}$, on a $xa \in \mathfrak{a}$.

Un sous-ensemble $\mathfrak{a} \subset A$ est un *idéal bilatère* si c'est un idéal à gauche et à droite, c'est-à-dire :

(1) $\mathfrak{a}$ est un sous-groupe additif de $(A, +)$;

(2) pour tout $a \in A$ et tout $x \in \mathfrak{a}$, on a $ax \in \mathfrak{a}$ et $xa \in \mathfrak{a}$.

Lorsque A est commutatif, ces trois notions coïncident. Dans ce cas, on parle seulement d'idéal de l'anneau A.

4.2. Exemples

(1) Les ensembles $\{0_A\}$ et A sont des idéaux bilatères de A.

(2) Pour tout $x \in A$, les ensembles

$$\{ax \mid a \in A\}, \quad \{xa \mid a \in A\}$$

sont respectivement des idéaux à gauche et à droite de A. De plus, l'ensemble

$$\left\{ \sum_{i=1}^{n} a_i x a'_i \mid n \geqslant 0,\ a_i,\ a'_i \in A \right\}$$

est un idéal bilatère de A.

(3) Tout idéal de $\mathbb{Z}$ est de la forme $n\mathbb{Z}$ $(= \mathbb{Z}n)$, $n \in \mathbb{Z}$ (on peut même choisir $n \geqslant 0$, et dans ce cas n est unique). En effet, un idéal étant un sous-groupe additif de $\mathbb{Z}$, il est nécessairement de cette forme par la proposition VI-2.4, et l'on vérifie aisément que c'est bien un idéal de $\mathbb{Z}$.

Le lemme suivant facilite les vérifications pour montrer qu'un sous-ensemble $\mathfrak{a}$ est un idéal.

4.3. Lemme. *Soit $\mathfrak{a}$ un sous-ensemble de A. Alors, $\mathfrak{a}$ est un idéal à gauche si, et seulement si, :*

(1) *$\mathfrak{a}$ est non vide ;*

(2) *pour tous x, $y \in \mathfrak{a}$ et tout $a \in A$, on a $x + ay \in \mathfrak{a}$.*

Démonstration. Le sens direct est clair. Supposons maintenant que $\mathfrak{a}$ satisfasse aux conditions ci-dessus. Pour tous x, $y \in A$. Alors, on a

$$x + (-1_A)y = x - y \in \mathfrak{a},$$

et donc $\mathfrak{a}$ est un sous-groupe additif de A (puisque $\mathfrak{a}$ est aussi non vide). En particulier, il contient 0_A, et donc pour tout $x \in \mathfrak{a}$ et tout $a \in A$, on a

$$ax = 0_A + ax \in \mathfrak{a}.$$

Ainsi, $\mathfrak{a}$ est un idéal à gauche. □

4.4. Remarque. On laisse le soin au lecteur d'énoncer et de démontrer des critères analogues pour les idéaux à droite et les idéaux bilatères.

Le lemme suivant est l'analogue du lemme VI-2.5.

4.5. Lemme. *Soit $f : A \longrightarrow B$ un morphisme d'anneaux, et soit $\mathfrak{b}$ un idéal à gauche (resp. à droite, resp. bilatère) de B. Alors, $f^{-1}(\mathfrak{b})$ est un idéal à gauche (resp. à droite, resp. bilatère) de A. En particulier,* $\mathrm{Ker}(f)$ *est un idéal bilatère de A.*

Démonstration. On sait déjà que $f^{-1}(\mathfrak{b})$ est un sous-groupe de A par le lemme VI-2.5. Soit $a \in A$ et soit $x \in f^{-1}(\mathfrak{b})$. Alors, on a $f(x) \in \mathfrak{b}$, et donc

$$f(ax) = f(a)f(x) \in \mathfrak{b},$$

car $\mathfrak{b}$ est un idéal à gauche de B. Ainsi, $ax \in f^{-1}(\mathfrak{b})$ et $f^{-1}(\mathfrak{b})$ est donc un idéal à gauche de A. On démontre de même que si $\mathfrak{b}$ est un idéal à droite (resp. bilatère), il en est de même de $f^{-1}(\mathfrak{b})$. Le dernier point est immédiat puisque $\{0_B\}$ est un idéal bilatère de B. □

Attention ! Si $\mathfrak{a}$ est un idéal (à gauche, à droite ou bilatère) de A, $f(\mathfrak{a})$ n'est pas nécessairement un idéal (à gauche, à droite ou bilatère) de B. On laisse le soin au lecteur de trouver un contre-exemple. Cela devient vrai néanmoins si f est surjectif.

4.6. Lemme. *L'intersection d'une famille d'idéaux à gauche (resp. à droite, resp. bilatères) de l'anneau A est un idéal à gauche (resp. à droite, resp. bilatère) de A.*

Démonstration. On démontre le résultat dans le cas des idéaux à gauche, les deux autres cas étant similaires.

Soit $(\mathfrak{a}_i)_{i\in I}$ une famille d'idéaux à gauche de A. On sait déjà que c'est un sous-groupe additif de A par le lemme VI-3.1. Si maintenant $x \in \mathfrak{a}_i$ pour tout $i \in I$ et $a \in A$, on a $ax \in \mathfrak{a}_i$ pour tout $i \in I$, car $\mathfrak{a}_i$ est un idéal à gauche, et donc ax appartient à l'intersection des $\mathfrak{a}_i$. □

4.7. Définition. Soit A un anneau, et soit S une partie de A (qui peut être éventuellement vide). *L'idéal à gauche, resp. à droite, resp. bilatère de A engendré par S*, noté $(S)_g$, resp. $(S)_d$, resp. (S), est l'intersection de tous les idéaux à gauche, resp. à droite, resp. bilatères de A contenant S. Autrement dit, $(S)_g$, resp. $(S)_d$, resp. (S), est le plus petit idéal à gauche, resp. à droite, resp. bilatère de A contenant S (au sens de l'inclusion).

L'idéal engendré jouit des mêmes propriétés que le sous-groupe engendré. Par exemple, $(S)_g = S$ si, et seulement si, S est un idéal à gauche de A.

Si $S = \{s_1, \ldots, s_m\}$, on note $(s_1, \ldots, s_m)_g$ plutôt que $(\{s_1, \ldots, s_m\})_g$, et de même pour les autres cas.

De même que dans le cas des sous-groupes, on a une description précise de l'idéal engendré par une partie.

4.8. Proposition. *Soit A un anneau, et soit S une partie de A. Alors, on a les égalités suivantes :*

(1) $(S)_g = \{a_1 s_1 + \cdots + a_n s_n \mid n \geqslant 0,\ s_i \in S,\ a_i \in A\}$;

(2) $(S)_d = \{s_1 a_1 + \cdots + s_n a_n \mid n \geqslant 0,\ s_i \in S,\ a_i \in A\}$;

(3) $(S) = \{a_1 s_1 a'_1 + \cdots + a_n s_n a'_n \mid n \geqslant 0,\ s_i \in S,\ a_i, a'_i \in A\}$.

Démonstration. Démontrons (1), les deux autres égalités s'établissant de manière similaire.

Montrons tout d'abord l'inclusion « $\supset$ ». Puisque $S \subset (S)_g$ par définition, pour tous $s_1, \ldots, s_n \in S$, et tous $a_1, \ldots, a_n \in A$, on a $a_1 s_1, \ldots, a_n s_n \in (S)_g$, car $(S)_g$ est un idéal à gauche. Puisque $(S)_g$ est aussi stable par addition,

on en déduit que $a_1s_1 + \cdots + a_ns_n \in (S)_g$. Pour montrer l'autre inclusion, on doit vérifier que l'ensemble

$$\{a_1s_1 + \cdots + a_ns_n \mid n \geqslant 0,\ s_i \in S,\ a_i \in A\}$$

est un idéal à gauche de A qui contient S, ce qui est facile en utilisant le lemme 4.3. □

4.9. Remarque. Pour tous $s_1, \ldots, s_m \in A$, on a les égalités suivantes :

$$\begin{aligned} (1)\quad (s_1, \ldots, s_m)_g &= \{a_1s_1 + \cdots + a_ms_m \mid a_i \in A\} \\ &= As_1 + \cdots + As_m; \end{aligned}$$

$$\begin{aligned} (2)\quad (s_1, \ldots, s_m)_d &= \{s_1a_1 + \cdots + s_ma_m \mid a_i \in A\} \\ &= s_1A + \cdots + s_mA. \end{aligned}$$

En particulier, pour tout $s \in A$, on a

(i) $(s)_g = \{as \mid a \in A\} = As$;
(ii) $(s)_d = \{sa \mid a \in A\} = sA$;

On a aussi

$$(s) = \Big\{ \sum_{i=1}^{n} a_isa'_i \mid n \geqslant 0,\ a_i, a'_i \in A \Big\}.$$

4.10. Remarque. Lorsque A est commutatif, les idéaux $(S)_g$, $(S)_d$ et (S) sont tous égaux. En particulier, si $a \in A$, on notera indifféremment (a), Aa ou aA l'idéal de A engendré par a.

Les idéaux pouvant être engendrés par un seul élément méritent un nom spécifique.

4.11. Définition. Un idéal $\mathfrak{a}$ à gauche (resp. à droite, resp. bilatère) d'un anneau A est dit *principal* s'il peut être engendré par un élément, c'est-à-dire s'il existe $a \in A$ tel que $\mathfrak{a} = (a)_g$ (resp. $(a)_d$, resp. (a)).

Un anneau A est *principal* s'il est intègre, et si tout idéal de A est principal.

4.12. Exemple. L'anneau $\mathbb{Z}$ est principal d'après l'exemple 4.2 (2).

4.13. Remarque. Nous verrons au cours de cet ouvrage des exemples d'anneaux principaux ou non principaux.

On continue ce paragraphe par un lemme élémentaire, mais fort utile.

4.14. Lemme. *Soit A un anneau, et soit $\mathfrak{a}$ un idéal à gauche (resp. à droite, resp. bilatère).*

(1) *On a $\mathfrak{a} = A$ si, et seulement si, $\mathfrak{a}$ contient un élément inversible de A.*

(2) *Si A est commutatif, pour tout $a \in A$, on a $(a) = A$ si, et seulement si, $a \in A^\times$.*

(3) *Si A est intègre, pour tous $a, b \in A$, on a $(a) = (b)$ si, et seulement si, il existe $u \in A^\times$ tel que $b = au$.*

Démonstration. Encore une fois, on démontre le premier point pour les idéaux à gauche. Si $\mathfrak{a} = A$, alors $\mathfrak{a}$ contient 1_A qui est inversible. Si maintenant $\mathfrak{a}$ contient $u \in A^\times$, alors pour tout $a \in A$, on a

$$a = (au^{-1})u \in \mathfrak{a},$$

car $\mathfrak{a}$ est un idéal à gauche, et donc $\mathfrak{a} = A$.

Supposons A commutatif, et soit $a \in A$. Si a est inversible, alors $(a) = A$ par le point précédent. Inversement, si $(a) = A$, alors $1_A \in A$, et donc il existe $b \in A$ tel que $ba = 1_A$. Ainsi, $a \in A^\times$ (car $ab = ba = 1_A$).

Supposons enfin que A soit intègre. Si $b = au$, $u \in A^\times$, on a $(b) \subset (a)$. Comme $u \in A^\times$, on a $a = u^{-1}b$ et donc $(a) \subset (b)$. Ainsi, on a $(a) = (b)$. Réciproquement, si $(a) = (b)$, il existe u, $v \in A$ tels que $b = au$ et $a = bv$, et donc $a = uva$. Si $a = 0_A$, alors $b = 0_A$, et donc $(a) = (b) = (0_A)$. Si $a \neq 0_A$, puisque A est intègre, on obtient $1_A = uv$ et donc $u \in A^\times$. □

On a alors le corollaire suivant.

4.15. Corollaire. *Soit A un anneau. Alors, A est un anneau à division si, et seulement si, A possède exactement deux idéaux à gauche (resp. à droite, resp. bilatères), à savoir (0_A) et A.*

Démonstration. Supposons que A soit un anneau à division. Alors, $0_A \neq 1_A$, et par suite (0_A) et A sont deux idéaux à gauches distincts. Si $\mathfrak{a}$ est un idéal à gauche non nul, il contient un élément non nul, donc inversible, et est donc égal à A par le point (1) du lemme précédent.

Inversement, supposons que (0_A) et A soient distincts, et soient les seuls idéaux à gauche. Alors, on a $0_A \neq 1_A$. Soit $a \in A \setminus \{0\}$. Alors, $(a)_g$ est un idéal à gauche non nul, donc égal à A. Il existe donc $b \in A$ tel que $ba = 1_A$. Remarquons que b est non nul, car $0_A \neq 1_A$. Par le même raisonnement, il existe $c \in A$ tel que $cb = 1_A$. On a alors

$$a = 1_A\, a = (cb)a = c(ba) = c\, 1_A = c.$$

Par conséquent, on a aussi $ab = 1_A$, et a est donc inversible. Ainsi, A est un anneau à division.

Le cas des idéaux à droite et bilatères se traitant de même, cela achève la démonstration. □

Nous achevons ce paragraphe en démontrant (entre autres) que l'ensemble des éléments nilpotents d'un anneau commutatif est un idéal.

4.16. Lemme. *Soit A un anneau commutatif. Alors :*

(1) *l'ensemble* $\mathrm{Nil}(A)$ *est un idéal de A ;*

(2) *pour tout $u \in A^\times$, et tout $a \in \mathrm{Nil}(A)$, on a $u \pm a \in A^\times$.*

Démonstration

(1) Remarquons que l'ensemble $\mathrm{Nil}(A)$ est non vide, puisqu'il contient 0_A. Soient a, $b \in \mathrm{Nil}(A)$, et soient $n, m \geqslant 1$ deux entiers tels que

$$a^n = 0_A \quad \text{et} \quad b^m = 0_A.$$

Soit $k \in [\![0, n+m]\!]$. Si $k \geqslant n$, on a

$$a^k = a^{k-n}a^n = 0_A,$$

et si $0 \leqslant k < n$, on a

$$b^{n+m-k} = b^{n-k}b^m = 0_A.$$

En particulier, on a

$$a^k b^{n+m-k} = 0_A, \quad \text{pour tout } k \in [\![0, n+m]\!].$$

La formule du binôme de Newton implique alors que $(a-b)^{n+m} = 0_A$. Ainsi, $a - b \in \mathrm{Nil}(A)$, et $\mathrm{Nil}(A)$ est un sous-groupe de A. Enfin, si $x \in A$, on a

$$(ax)^n = a^n x^n = 0_A,$$

et donc $ax \in A$. Cela démontre que $\mathrm{Nil}(A)$ est un idéal de A.

(2) Soit $a \in \mathrm{Nil}(A)$ et soit $u \in A^\times$. Puisque $\mathrm{Nil}(A)$ est un idéal de A par (1), on a $u^{-1}a \in \mathrm{Nil}(A)$, et d'après le lemme 2.11, $1_A \pm u^{-1}a \in A^\times$. Mais alors, $u \pm a = u(1_A \pm u^{-1}a) \in A^\times$, d'où le résultat. □

Ce résultat devient faux si A n'est pas commutatif, comme le montre l'exemple suivant.

4.17. Exemple. Soit $A = \mathrm{M}_2(\mathbb{C})$. Alors, les éléments

$$N_1 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{et} \quad N_2 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}$$

sont nilpotents, mais la matrice

$$N_1 + N_2 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

n'est pas nilpotente. De plus, la matrice $U = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$ est inversible, mais

la matrice
$$U + N_1 = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$
est de déterminant nul, donc non inversible.

Dorénavant, on écrira 0 et 1, plutôt que 0_A et 1_A, afin de ne pas alourdir les notations.

5. Produit direct d'anneaux

Nous allons maintenant nous intéresser au problème suivant : à quelles conditions un anneau A donné est-il isomorphe à un produit direct d'anneaux.

On commence par un lemme, qui démontre que les idempotents interviennent naturellement dans cette question.

5.1. Lemme. *Soient $A, A_1, \ldots, A_n$ des anneaux. On suppose qu'il existe un isomorphisme d'anneaux*
$$\varphi : A \xrightarrow{\sim} A_1 \times \cdots \times A_n.$$
Pour tout $i \in [\![1, n]\!]$, on pose
$$e_i = \varphi^{-1}(0, \ldots, 0, 1, 0, \ldots, 0),$$
et l'on note $\pi_i : A_1 \times \cdots \times A_n \longrightarrow A_i$ la projection sur la i-ème composante. Alors, on a les propriétés suivantes :

(1) *pour tout $i \in [\![1, n]\!]$, e_i est un idempotent, et $e_i a = a e_i$, pour tout $a \in A$;*
(2) *pour tous $i, j \in [\![1, n]\!]$, $i \neq j$, on a $e_i e_j = e_j e_i = 0$;*
(3) *on a $1 = e_1 + \cdots + e_n$;*
(4) *pour tout $i \in [\![1, n]\!]$, le morphisme d'anneaux $\pi_i \circ \varphi : A \longrightarrow A_i$ induit par restriction un isomorphisme d'anneaux $\varphi_i : e_i A \simeq A_i$. De plus, pour tout $a \in A$, on a*
$$\varphi_i(e_i a) = \pi_i\big(\varphi(a)\big).$$

Démonstration. Pour tout $i \in [\![1, n]\!]$, posons $\varepsilon_i = (0, \ldots, 0, 1, 0, \ldots, 0)$. On vérifie alors aisément que l'on a les propriétés suivantes.

(1) Pour tout $i \in [\![1, n]\!]$, ε_i est un idempotent, et l'on a
$$\varepsilon_i(a_1, \ldots, a_n) = (a_1, \ldots, a_n)\varepsilon_i = (0, \ldots, 0, a_i, 0, \ldots, 0)$$
pour tout $(a_1, \ldots, a_n) \in A_1 \times \cdots \times A_n$.
(2) Pour tous $i, j \in [\![1, n]\!]$, $i \neq j$, on a $\varepsilon_i \varepsilon_j = \varepsilon_j \varepsilon_i = (0, \ldots, 0)$.

(3) On a $1_{A_1\times\cdots\times A_n} = (1,\dots,1) = \varepsilon_1 + \cdots + \varepsilon_n$.

On en déduit alors les points (1) – (3) de l'énoncé en appliquant φ^{-1} à ces égalités, et en utilisant le fait que φ^{-1} est un isomorphisme d'anneaux. Il reste à établir (4) et (5). Puisque e_i commute avec tout élément de A, on a

$$e_i A e_i = e_i^2 A = e_i A.$$

D'après le lemme 2.15, e_iA est alors un anneau, de neutre multiplicatif e_i. Comme $\pi_i \circ \varphi$ est un morphisme d'anneaux, l'application φ_i respecte la somme et le produit. Il reste à voir que l'image de l'élément neutre de l'anneau e_iA, à savoir e_i, est bien 1. Mais on a

$$\varphi_i(e_i) = \pi_i\big(\varphi(e_i)\big) = \pi_i(\varepsilon_i) = 1,$$

et φ_i est bien un morphisme d'anneaux.

Soit $a \in A$. Remarquons que $\pi_i\big(\varphi(a)\big)$ est la i-ème composante de $\varphi(a)$. Mais alors, on a

$$\begin{aligned}\varphi(e_i a) &= \varphi(e_i)\varphi(a)\\ &= \varepsilon_i \varphi(a)\\ &= (0,\dots,0,1,0,\dots,0)\big(\pi_1(\varphi(a)),\dots,\pi_n(\varphi(a))\big)\\ &= \big(0,\dots,0,\pi_i\big(\varphi(a)\big),0,\dots,0\big).\end{aligned}$$

En particulier,

$$\varphi_i(e_i a) = \pi_i\big(\varphi(a)\big), \quad \text{pour tout } a \in A.$$

Comme $\pi_i \circ \varphi$ est surjective, cela montre la surjectivité de φ_i. Il reste donc à montrer l'injectivité de φ_i. Mais si $\varphi_i(e_i a) = \pi_i\big(\varphi(a)\big) = 0$, les considérations précédentes montrent que $\varphi(e_i a) = (0,\dots,0)$. Par injectivité de φ, on a $e_i a = 0$, d'où le résultat. □

5.2. Remarque. L'idempotent e_i de l'énoncé est non nul si, et seulement si, $1 \neq 0$ dans A_i, c'est-à-dire si, et seulement si, A_i est non trivial.

Ce lemme conduit naturellement à proposer la définition suivante.

5.3. Définition. Soit A un anneau. Un idempotent $e \in A$ est dit *central* si l'on a $e \in Z(A)$, c'est-à-dire s'il commute avec tout élément de A.

Deux idempotents e, $f \in A$ sont *orthogonaux* si $ef = fe = 0$.

Enfin, un *système fondamental d'idempotents orthogonaux* est une famille finie d'idempotents **tous non nuls** $(e_1,\dots,e_n)$ vérifiant les conditions suivantes :

(1) pour tout $i \in [\![1,n]\!]$, e_i est un idempotent central ;

(2) pour tous i, $j \in [\![1,n]\!]$, $i \neq j$, e_i et e_j sont orthogonaux ;

(3) on a $1 = e_1 + \cdots + e_n$.

5.4. Exemples

(1) Si $e \in A$ est un idempotent central non trivial, alors $(e, 1-e)$ est un système fondamental d'idempotents orthogonaux.

(2) Soient $A_1, \ldots, A_n$ des anneaux non triviaux. Pour tout $i \in [\![1, n]\!]$, posons $\varepsilon_i = (0, \ldots, 0, 1, 0, \ldots, 0) \in A_1 \times \cdots \times A_n$. Alors, $(\varepsilon_1, \ldots, \varepsilon_n)$ est un système fondamental d'idempotents orthogonaux de $A_1 \times \cdots \times A_n$.

On continue par un résultat qui permet de savoir lorsqu'un anneau se décompose en produit direct d'anneaux.

5.5. Théorème. *Soit A un anneau. Alors, les propriétés suivantes sont équivalentes :*

(1) *il existe des anneaux non triviaux $A_1, \ldots, A_n$ tels que*

$$A \simeq A_1 \times \cdots \times A_n ;$$

(2) *l'anneau A possède un système fondamental de n idempotents orthogonaux $(e_1, \ldots, e_n)$.*

Dans ce cas, l'application

$$\begin{aligned} \theta : A &\longrightarrow e_1 A \times \cdots \times e_n A \\ a &\longmapsto (e_1 a, \ldots, e_n a) \end{aligned}$$

est un isomorphisme d'anneaux, d'inverse

$$\begin{aligned} \theta^{-1} : e_1 A \times \cdots \times e_n A &\longrightarrow A \\ (a_1, \ldots, a_n) &\longmapsto \sum_{i=1}^{n} e_i a_i . \end{aligned}$$

Démonstration. Le sens (1) $\Longrightarrow$ (2) est donné par le lemme 5.1 et la remarque qui suit. Supposons maintenant que $(e_1, \ldots, e_n)$ soit un système fondamental d'idempotents orthogonaux de A. Par le lemme 2.15, et puisque e_i est un idempotent central, l'ensemble $e_i A e_i = e_i^2 A = e_i A$ est un anneau, de neutre additif 0, et de neutre multiplicatif e_i. En particulier, puisque $e_i \neq 0$, l'anneau $e_i A$ est non trivial.

On vérifie alors sans mal que θ est un morphisme d'anneaux. Soit $a \in A$ tel que $e_i a = 0$ pour tout $i \in [\![1, n]\!]$. Alors, on a

$$a = 1 \cdot a = (e_1 + \cdots + e_n) a = e_1 a + \cdots + e_n a = 0.$$

Ainsi, θ est injective. Remarquons maintenant que, pour tout $a \in A$, et tout entier $i \in [\![1, n]\!]$, on a

$$\begin{aligned} \theta(e_i a) &= (e_1 e_i a, \ldots, e_n e_i a) \\ &= (0, \ldots, 0, e_i^2 a, 0, \ldots, 0) \\ &= (0, \ldots, 0, e_i a, 0, \ldots, 0), \end{aligned}$$

car pour tout $j \neq i$, on a $e_j e_i = 0$.

Mais alors, pour tout $(e_1a_1, \ldots, e_na_n) \in e_1A \times \cdots \times e_nA$, on a

$$\theta\Big(\sum_{i=1}^{n} e_ia_i\Big) = \sum_{i=1}^{n} \theta(e_ia_i) = (e_1a_1, \ldots, e_na_n),$$

d'où la surjectivité de θ. Cette égalité montre également que l'isomorphisme inverse est bien celui de l'énoncé. Cela achève la démonstration. □

5.6. Remarque. Soient $A, A_1, \ldots, A_n$ des anneaux non triviaux. On suppose qu'il existe un isomorphisme d'anneaux

$$\varphi : A \xrightarrow{\sim} A_1 \times \cdots \times A_n.$$

Pour tout $i \in [\![1, n]\!]$, on pose

$$e_i = \varphi^{-1}\big((0, \ldots, 0, 1, 0, \ldots, 0)\big),$$

et l'on note $\pi_i : A_1 \times \cdots \times A_n \longrightarrow A_i$ la projection sur la i-ème composante. Alors, l'isomorphisme inverse de φ est donné par

$$\begin{aligned} \varphi^{-1} : A_1 \times \cdots \times A_n &\longrightarrow A \\ (a_1, \ldots, a_n) &\longmapsto \sum_{i=1}^{n} e_i\tilde{a}_i, \end{aligned}$$

où $\tilde{a}_i$ est un antécédent arbitraire de a_i par le morphisme surjectif

$$\pi_i \circ \varphi : A \longrightarrow A_i.$$

En effet, rappelons tout d'abord que, pour tout $i \in [\![1, n]\!]$, on a un isomorphisme d'anneaux

$$\begin{aligned} \varphi_i : e_iA &\longrightarrow A_i \\ e_ia &\longmapsto \pi_i\big(\varphi(a)\big), \end{aligned}$$

et cela d'après le lemme 5.1 (4).

On a donc un isomorphisme

$$\begin{aligned} \rho : e_1A \times \cdots \times e_nA &\longrightarrow A_1 \times \cdots \times A_n \\ (e_1a_1, \ldots, e_na_n) &\longmapsto \big(\pi_1\big(\varphi(a_1)\big), \ldots, \pi_n\big(\varphi(a_n)\big)\big). \end{aligned}$$

Mais alors, pour tout $a \in A$, on a

$$\rho\big(\theta(a)\big) = \big(\pi_1\big(\varphi(a)\big), \ldots, \pi_n\big(\varphi(a)\big)\big) = \varphi(a).$$

Ainsi, $\varphi = \rho \circ \theta$. Soit maintenant $(a_1, \ldots, a_n) \in A_1 \times \cdots \times A_n$. Si $\tilde{a}_i \in A$ est un antécédent arbitraire de a_i par $\pi_i \circ \varphi$, on a donc

$$\rho^{-1}\big((a_1, \ldots, a_n)\big) = (e_1\tilde{a}_1, \ldots, e_n\tilde{a}_n),$$

et par conséquent

$$\theta^{-1}\big(\rho^{-1}\big((a_1,\dots,a_n)\big)\big) = \sum_{i=1}^{n} e_i(e_i\tilde{a}_i) = \sum_{i=1}^{n} e_i^2\tilde{a}_i = \sum_{i=1}^{n} e_i\tilde{a}_i.$$

Par suite, on a bien

$$\varphi^{-1}\big((a_1,\dots,a_n)\big) = \sum_{i=1}^{n} e_i\tilde{a}_i.$$

5.7. Corollaire. *Soit A un anneau. Alors, les propriétés suivantes sont équivalentes :*

(1) *il existe des anneaux A_1, A_2 non triviaux tels que $A \simeq A_1 \times A_2$;*

(2) *l'anneau A possède un idempotent central non trivial.*

5.8. Remarques

(1) Dans tous les résultats précédents, le fait que les idempotents soient centraux est crucial. Par exemple, il existe des anneaux possédant des idempotents non triviaux, mais qui ne sont pas isomorphes à un produit direct de deux anneaux non triviaux (cf. exercice 31).

(2) On peut se passer de la condition de non-nullité des idempotents dans la définition d'un système fondamental d'idempotents orthogonaux. Dans ce cas, le théorème 5.5 reste vrai si l'on autorise les anneaux A_i à être triviaux. Néanmoins, cela ne présente qu'un intérêt négligeable.

Le prochain paragraphe est consacré à l'étude d'un exemple, à savoir l'anneau des polynômes.

6. Anneaux de polynômes

Dans ce paragraphe, on définit l'anneau des polynômes à coefficients dans un anneau A, et l'on étudie quelques unes de ses propriétés.

Soit A un anneau, et soit $\mathcal{S}$ le groupe additif des suites $(a_n)_{n\geqslant 0}$ d'éléments de A (pour l'addition terme à terme).

Pour tout $n \geqslant 0$, on note $e_n \in \mathcal{S}$ la suite dont tous les termes sont nuls, sauf celui d'indice n, qui vaut 1.

Si $s = (a_n)_{n\geqslant 0}$, $s' = (b_n)_{n\geqslant 0} \in \mathcal{S}$, on pose

$$ss' = \Big(\sum_{k=0}^{n} a_k b_{n-k}\Big)_{n\geqslant 0} = \Big(\sum_{\substack{i,j\geqslant 0\\ i+j=n}} a_i b_j\Big)_{n\geqslant 0}.$$

6.1. Proposition. *L'ensemble $\mathcal{S}$ muni de l'addition et de la loi de multiplication précédente est un anneau, de neutre multiplicatif e_0.*

Démonstration. Il reste en fait à montrer les propriétés relatives à la multiplication, puisque l'on sait déjà que $(\mathcal{S}, +)$ est un groupe additif. Commençons par l'associativité. Si $s = (a_n)_{n\geqslant 0}$, $s' = (b_n)_{n\geqslant 0}$, $s'' = (c_n)_{n\geqslant 0} \in \mathcal{S}$, on a par définition

$$(ss')s'' = \Big(\sum_{\ell=0}^{n}\Big(\sum_{k=0}^{\ell} a_k b_{\ell-k}\Big)c_{n-\ell}\Big)_{n\geqslant 0} = \Big(\sum_{i_1+i_2+i_3=n}(a_{i_1}b_{i_2})c_{i_3}\Big)_{n\geqslant 0}.$$

De même, on a

$$s(s's'') = \Big(\sum_{i_1+i_2+i_3=n} a_{i_1}(b_{i_2}c_{i_3})\Big)_{n\geqslant 0}.$$

Comme la multiplication de A est associative, il en est donc de même de celle de $\mathcal{S}$.

La distributivité et la commutativité proviennent aisément de la définition. Enfin, on vérifie que e_0 est un élément neutre pour la multiplication. □

On introduit maintenant de nouvelles notations. Si $a \in A$ et $(a_n) \in \mathcal{S}$, on pose

$$a(a_n)_{n\geqslant 0} = (aa_n)_{n\geqslant 0}.$$

C'est tout simplement le produit de $(a, 0, 0, \ldots)$ par $(a_n)_{n\geqslant 0}$.

Avec cette nouvelle notation, la suite $(0, \ldots, 0, a_n, 0, \ldots, 0)$ est donc $a_n e_n$. Nous allons donc être un peu cavalier, et écrire un élément de $\mathcal{S}$ sous la forme $\sum_{n\geqslant 0} a_n e_n$.

Notons que cette notation est purement formelle, la somme précédente ayant a priori un nombre infini de termes. Si l'on veut donner à cette somme un vrai sens mathématique, il faudrait introduire une topologie, ce que l'on fait dans le chapitre XXIII. Cette commodité d'écriture va néanmoins nous permettre de visualiser un peu mieux les lois définies plus haut. On vérifie facilement en utilisant la définition que l'on a

$$e_1^n = e_n, \quad \text{pour tout } n \geqslant 0.$$

Posons $X = e_1$ et notons $A[[X]]$ l'anneau $\mathcal{S}$. D'après ce qui précède un élément $S \in A[[X]]$ s'écrit

$$S = \sum_{n\geqslant 0} a_n X^n.$$

De plus, les opérations sur $A[[X]]$ sont définies par

$$\sum_{n\geqslant 0} a_n X^n + \sum_{n\geqslant 0} b_n X^n = \sum_{n\geqslant 0}(a_n + b_n)X^n$$

et

$$\Big(\sum_{n\geqslant 0} a_n X^n\Big)\Big(\sum_{n\geqslant 0} b_n X^n\Big) = \sum_{n\geqslant 0}\Big(\sum_{k=0}^{n} a_k b_{n-k}\Big)X^n.$$

Par définition, on a $\sum_{n\geqslant 0} a_n X^n = 0$ si, et seulement si, $a_n = 0$ pour tout entier $n \geqslant 0$.

Le neutre multiplicatif est donné par X^0, que l'on notera 1, et le neutre additif est donné par $\sum_{n\geqslant 0} 0\ X^n$, que l'on notera 0.

6.2. Définition. Soit A un anneau. L'anneau $A[[X]]$ est appelé l'anneau des *séries formelles* à coefficients dans A.

L'anneau des séries formelles sera étudié en détail dans un chapitre ultérieur. Ici, nous allons plutôt nous intéresser à un sous-anneau particulier de $A[[X]]$, à savoir l'anneau des polynômes à coefficients dans A.

6.3. Lemme. *Soit A un anneau. Alors, l'ensemble $A[X]$ des séries formelles $\sum_{n\geqslant 0} a_n X^n$, où les coefficients $a_n \in A$ sont tous nuls sauf un nombre fini, est un sous-anneau de $A[[X]]$.*

Démonstration. Il est clair que les séries formelles 0 et 1 sont des éléments de $A[X]$, et que $A[X]$ est stable par addition et par passage à l'opposé. Soient $\sum_{n\geqslant 0} a_n X^n$, $\sum_{n\geqslant 0} b_n X^n \in A[X]$. Il existe donc N, $N' \geqslant 0$ tels que

$$a_n = 0, \text{ pour tout } n \geqslant N \quad \text{et} \quad b_n = 0, \text{ pour tout } n \geqslant N'.$$

Mais alors pour tout $n \geqslant N + N'$, on a $\sum_{k=0}^{n} a_k b_{n-k} = 0$, car soit $k \geqslant N$, et dans ce cas $a_k = 0$, soit $k < N$, mais dans ce cas

$$n - k \geqslant N + N' - k > N',$$

et $b_{n-k} = 0$. Ainsi, $A[X]$ est aussi stable par produit, d'où le résultat souhaité. □

6.4. Remarque. On montre aisément que $A[[X]]$ est commutatif si, et seulement si, A l'est.

On vérifie que l'application

$$\begin{aligned} A &\longrightarrow A[[X]] \\ a &\longmapsto aX^0 \end{aligned}$$

est un morphisme d'anneaux injectif, qui identifie l'anneau A à un sous-anneau de $A[[X]]$, et même à un sous-anneau de $A[X]$. On identifiera A avec son image dans $A[X]$ et dans $A[[X]]$, et l'on notera a plutôt que $a\ 1$ ou aX^0.

Les éléments de A, vus comme polynômes (resp. des séries formelles), seront appelés des polynômes *constants* (resp. des séries formelles *constantes*).

6.5. Définition. Soit A un anneau. L'anneau $A[X]$ est appelé anneau des *polynômes* à coefficients dans A.

6.6. Remarques

(1) Notons que, dans le cas de l'anneau $A[X]$, la notation sommatoire a bien un sens, puisque l'on somme un nombre fini de termes.

(2) Encore une fois, $A[X]$ est commutatif si, et seulement si, A l'est.

(3) Pour tout $n \geqslant 0$, on définit l'anneau $A[X_1, \ldots, X_n]$ par récurrence en posant $A[X_1, \ldots, X_n] = A$ pour $n = 0$, et

$$A[X_1, \ldots, X_n] = A[X_1, \ldots, X_{n-1}][X_n],$$

pour tout $n \geqslant 1$.

Nous allons maintenant illustrer les notions introduites dans ce chapitre dans le cas de l'anneau $A[X]$. On commence par une définition.

6.7. Définition. Soit A un anneau, et soit $f \in A[X] \setminus \{0\}$. Puisque $f \neq 0$, on peut écrire

$$f = a_d X^d + a_{d-1} X^{d-1} + \cdots + a_0,$$

avec $a_i \in A$, et $a_d \in A \setminus \{0\}$. L'entier $d \geqslant 0$ est appelé le *degré* de f, et est noté $\deg(f)$. Par convention, on pose $\deg(0) = -\infty$. On a donc $\deg(f) \geqslant 0$ si, et seulement si, f est non nul.

Le coefficient a_d est appelé le *coefficient dominant* de f. On dit qu'un polynôme est *unitaire* si son coefficient dominant est égal à 1.

Le lemme suivant est souvent utile dans les calculs.

6.8. Lemme. *Soit A un anneau non trivial, et soit $g \in A[X]$ un polynôme non nul. On suppose que le coefficient dominant de g n'est pas un diviseur de zéro dans A.*
Alors, pour tout $f \in A[X]$, on a

$$\deg(fg) = \deg(f) + \deg(g).$$

En particulier, l'application

$$\begin{aligned} r_g \colon A[X] &\longrightarrow A[X] \\ f &\longmapsto fg \end{aligned}$$

est injective.

Démonstration. Si f est nul, l'égalité désirée est triviale. Supposons maintenant que f soit non nul, et écrivons

$$f = a_n X^n + \cdots + a_1 X + a_0, \quad g = b_m X^m + \cdots + b_1 X + b_0, \quad a_n, b_m \in A \setminus \{0\}.$$

En utilisant la définition, on voit que l'on a

$$fg = a_n b_m X^{n+m} + \text{ des termes en } X^i, \quad i < n + m.$$

Comme b_m n'est pas un diviseur de zéro et $a_n \neq 0$, on a $a_n b_m \neq 0$, ce qui démontre l'égalité

$$\deg(fg) = n + m = \deg(f) + \deg(g).$$

Cela démontre aussi que si $f \neq 0$, on a $fg \neq 0$.

Remarquons maintenant que r_g est un morphisme de groupes additifs. Or, le point précédent montre que $\mathrm{Ker}(r_g) = \{0\}$. Par conséquent, r_g est injective. □

Contrairement au cas de l'anneau des polynômes à coefficients dans un corps, l'anneau $A[X]$ n'est pas nécessairement intègre. Le lemme suivant donne des conditions nécessaires et suffisantes pour cela soit le cas.

6.9. Lemme. *Soit A un anneau commutatif non trivial. Alors, les propriétés suivantes sont équivalentes :*

(1) *l'anneau $A[X]$ est intègre ;*

(2) *l'anneau A est intègre ;*

(3) *pour tous $f, g \in A[X]$, on a $\deg(fg) = \deg(f) + \deg(g)$.*

Démonstration

(1) $\Longrightarrow$ (2). C'est immédiat, car un sous-anneau d'un anneau intègre est intègre.

(2) $\Longrightarrow$ (3). Si g est nul, l'égalité est triviale. Si g est non nul, son coefficient dominant est non nul, et n'est donc pas un diviseur de zéro puisque A est intègre. Le résultat découle alors du lemme 6.8.

(3) $\Longrightarrow$ (1). Puisque l'anneau A est commutatif et non trivial, il en est de même de l'anneau $A[X]$.

Soient f et g deux polynômes non nuls. On a donc $\deg(f) \geqslant 0$ et $\deg(g) \geqslant 0$. Mais alors, on a

$$\deg(fg) = \deg(f) + \deg(g) \geqslant 0,$$

ce qui entraîne que fg est non nul, d'où l'intégrité de $A[X]$. □

On continue ce paragraphe en expliquant comment diviser deux polynômes.

6.10. Théorème (Division de polynômes généralisée). *Soit A un anneau non trivial. Soient $f, g \in A[X]$, avec $g \neq 0$. On suppose que le coefficient dominant de g est un élément inversible de A. Alors, il existe deux polynômes $Q, R \in A[X]$, avec $\deg(R) < \deg(g)$ tels que $f = Qg + R$. De plus, Q et R sont uniques.*

Démonstration. Si $\deg(g) > \deg(f)$, on pose $Q = 0$ et $R = f$. On suppose maintenant que $\deg(g) \leqslant \deg(f)$. On va démontrer l'existence de Q et R par récurrence sur $\deg(f)$. Soit (H_n) la propriété :

(H_n) : Pour tous $f, g \in A[X], g \neq 0$ tels que $0 \leqslant \deg(g) \leqslant \deg(f) \leqslant n$, et tels que le coefficient dominant de g soit inversible, il existe $Q, R \in A[X]$, tels que $\deg(R) < \deg(g)$ et $f = Qg + R$.

Si f est le polynôme nul, c'est clair, car on prend $Q = R = 0$. Si $n = 0$, alors $\deg(f) = \deg(g) = 0$. On a donc $f = a,\ g = b,\ a, b \in A^\times$. Dans ce cas, on pose $Q = ab^{-1}$ et $R = 0$. Ainsi, (H_0) est vraie.

Supposons (H_n) vraie pour un certain entier $n \geqslant 0$, et montrons (H_{n+1}).

Si $\deg(f) \leqslant n$, alors Q et R existent, car (H_n) est vraie, donc on peut supposer que $\deg(f) = n + 1$. Écrivons

$$f = a_{n+1}X^{n+1} + \cdots + a_1X + a_0, \quad g = b_mX^m + \cdots + b_1X + b_0.$$

Par hypothèse, on a $n + 1 \geqslant m$.

Soit $Q_1 = a_{n+1}b_m^{-1}X^{n+1-m}$ et $R_1 = f - a_{n+1}b_m^{-1}X^{n+1-m}g$. Remarquons que b_m^{-1} a un sens dans A, puisque $b_m \in A^\times$. Puisque $f, g \in A[X]$, on a aussi $Q_1, R_1 \in A[X]$. Par construction, le coefficient en X^{n+1} de R_1 est 0, donc $\deg(R_1) \leqslant n$.

Si $\deg(R_1) < \deg(g)$, on a fini, car on peut écrire

$$f = a_{n+1}b_m^{-1}X^{n+1-m}g + R_1,$$

et donc poser $R = R_1$ et $Q = a_{n+1}b_m^{-1}X^{n+1-m}$.

Supposons que $\deg(R_1) \geqslant \deg(g)$. Comme $\deg(g) \leqslant \deg(R_1) \leqslant n$, on peut appliquer (H_n), et donc il existe $Q_2, R_2 \in A[X]$ tels que $R_1 = Q_2g + R_2$, avec $\deg(R_2) < \deg(g)$.

Ainsi, on a $f = (a_{n+1}b_m^{-1}X^{n+1-m} + Q_2)g + R_2$, et $\deg(R_2) < \deg(g)$. Maintenant, il suffit de poser $Q = a_{n+1}b_m^{-1}X^{n+1-m} + Q_2$ et $R = R_2$.

Par conséquent, (H_{n+1}) est vraie, ce qui achève la récurrence.

Montrons maintenant l'unicité de Q et R. Supposons que l'on ait

$$f = Q_1g + R_1 = Q_2g + R_2,$$

avec $Q_i,\ R_i \in A[X]$ et $\deg(R_i) < \deg(g)$. Alors, on a $(Q_1 - Q_2)g = R_2 - R_1$.

Puisque le coefficient dominant de g est inversible, ce n'est pas un diviseur de zéro. Par le lemme 6.8, on a donc

$$\deg\big((Q_1 - Q_2)g\big) = \deg(Q_1 - Q_2) + \deg(g).$$

Si $Q_1 - Q_2 \neq 0$, on a donc

$$\deg(R_2 - R_1) = \deg(Q_1 - Q_2) + \deg(g) \geqslant \deg(g).$$

Néanmoins, il est facile de constater que $\deg(R_2 - R_1) < \deg(g)$, d'où une contradiction. Ainsi, $Q_1 - Q_2 = 0$, puis $R_2 - R_1 = 0$, d'où l'unicité. □

6.11. Définition. Les polynômes Q et R s'appellent respectivement le *quotient* et le *reste* de la division de f par g.

En pratique, la meilleure façon de calculer Q et R est de procéder comme dans le cas où $A = \mathbb{R}$, $\mathbb{C}$ (en posant la division).

6.12. Remarque. Notons que, dans l'énoncé du théorème 6.10, nous n'avons pas supposé l'anneau commutatif.

D'ailleurs, en remplaçant « $a_{n+1}b_m^{-1}$ » par « $b_m^{-1}a_{n+1}$ », on démontre aussi l'existence et l'unicité de polynômes $\widetilde{Q}, \widetilde{R} \in A[X]$ tels que

$$f = g\widetilde{Q} + \widetilde{R} \quad \text{et} \quad \deg(\widetilde{R}) < \deg(g).$$

Attention ! Le théorème 6.10 n'est pas vrai si le coefficient dominant de g n'est pas inversible.

Par exemple, X, $2X \in \mathbb{Z}[X]$, mais le quotient de X par $2X$ est $\frac{1}{2}$, qui n'est pas dans $\mathbb{Z}[X]$.

6.13. Remarque. Si K est un corps (et même un anneau à division), on peut donc diviser n'importe quel polynôme par un polynôme non nul.

Comme dans le cas de l'anneau $\mathbb{Z}$, l'existence d'une division euclidienne dans $K[X]$ a des conséquences intéressantes sur les idéaux.

6.14. Lemme. *Soit K un corps. Alors, $K[X]$ est un anneau principal.*

Démonstration. Puisqu'un corps K est un anneau intègre, il en est de même de $K[X]$ par le lemme 6.9. Soit $\mathfrak{a}$ un idéal de $K[X]$. Montrons qu'il est principal. Si $\mathfrak{a} = (0)$, il n'y a rien à faire. Supposons donc que $\mathfrak{a} \neq (0)$, et soit f_0 un élément de $\mathfrak{a} \setminus \{0\}$ de degré minimal. Montrons que $\mathfrak{a} = (f_0)$. Puisque $f_0 \in \mathfrak{a}$, on a $(f_0) \subset \mathfrak{a}$. Si maintenant $f \in \mathfrak{a}$, il existe $Q, R \in K[X]$ tels que $f = Qf_0 + R$ avec $\deg(R) < \deg(f_0)$ (voir la remarque 6.13). Puisque f, $f_0 \in \mathfrak{a}$, on a $R = f - Qf_0 \in \mathfrak{a}$, car $\mathfrak{a}$ est un idéal. Si R était non nul, cela contredirait la minimalité du degré de f_0. On a donc $R = 0$, soit encore $f = Qf_0 \in (f_0)$. On a donc finalement l'inclusion manquante, ce qui achève la démonstration. □

6.15. Remarque. La description des idéaux de $\mathbb{Z}$ et $K[X]$ s'inscrit en fait dans un contexte plus général, à savoir celui des anneaux euclidiens.

On s'intéresse maintenant aux morphismes d'anneaux de $A[X]$ vers un anneau quelconque B.

6.16. Proposition. *Soit $u : A \to B$ un morphisme d'anneaux commutatifs, et soit $b \in B$. Alors, il existe un unique morphisme d'anneaux*

$$\psi_{u,b} : A[X] \to B$$

tel que

$$\psi_{u,b}(a) = u(a), \quad \textit{pour tout } a \in A \quad \textit{et} \quad \psi_{u,b}(X) = b.$$

Il est défini par

$$\psi_{u,b}\Big(\sum_{n\geqslant 0} a_n X^n\Big) = \sum_{n\geqslant 0} u(a_n) b^n.$$

De plus, pour tout morphisme d'anneaux $\varphi : A[X] \longrightarrow B$, il existe un unique morphisme d'anneaux $u : A \longrightarrow B$ et un unique élément $b \in B$ tel que $\varphi = \psi_{u,b}$.

Démonstration. L'application $\psi_{u,b}$ satisfait clairement les conditions requises. Il reste à vérifier que c'est bien un morphisme d'anneaux. Pour commencer, on a $\psi_{u,b}(1) = u(1) = 1$. De plus,

$$\psi_{u,b}\Big(\sum_{n\geqslant 0} a_n X^n + \sum_{n\geqslant 0} b_n X^n\Big) = \psi_{u,b}\Big(\sum_{n\geqslant 0} (a_n + b_n) X^n\Big) = \sum_{n\geqslant 0} u(a_n + b_n) b^n.$$

Puisque u est un morphisme d'anneaux, on obtient

$$\begin{aligned} \psi_{u,b}\Big(\sum_{n\geqslant 0} a_n X^n + \sum_{n\geqslant 0} b_n X^n\Big) &= \sum_{n\geqslant 0} \big(u(a_n) + u(b_n)\big) b^n \\ &= \sum_{n\geqslant 0} u(a_n) b^n + \sum_{n\geqslant 0} u(b_n) b^n, \end{aligned}$$

et donc

$$\psi_{u,b}\Big(\sum_{n\geqslant 0} a_n X^n + \sum_{n\geqslant 0} b_n X^n\Big) = \psi_{u,b}\Big(\sum_{n\geqslant 0} a_n X^n\Big) + \psi_{u,b}\Big(\sum_{n\geqslant 0} b_n X^n\Big).$$

Finalement, on a

$$\begin{aligned} \psi_{u,b}\Big(\Big(\sum_{n\geqslant 0} a_n X^n\Big)\Big(\sum_{n\geqslant 0} b_n X^n\Big)\Big) &= \psi_{u,b}\Big(\sum_{n\geqslant 0}\Big(\sum_{j+k=n} a_j b_k\Big) X^n\Big) \\ &= \sum_{n\geqslant 0} u\Big(\sum_{j+k=n} a_j b_k\Big) b^n. \end{aligned}$$

Puisque u est un morphisme d'anneaux, on a

$$\begin{aligned}\psi_{u,b}\Big(\Big(\sum_{n\geqslant 0} a_n X^n\Big)\Big(\sum_{n\geqslant 0} b_n X^n\Big)\Big) &= \sum_{n\geqslant 0}\sum_{j+k=n} u(a_j)u(b_k)b^n \\ &= \sum_{n\geqslant 0}\sum_{j+k=n} u(a_j)u(b_k)b^j b^k \\ &= \sum_{n\geqslant 0}\sum_{j+k=n} \big(u(a_j)b^j\big)\big(u(b_k)b^k\big) \\ &= \Big(\sum_{n\geqslant 0} u(a_n)b^n\Big)\Big(\sum_{n\geqslant 0} u(b_n)b^n\Big),\end{aligned}$$

et ainsi

$$\psi_{u,b}\Big(\Big(\sum_{n\geqslant 0} a_n X^n\Big)\Big(\sum_{n\geqslant 0} b_n X^n\Big)\Big) = \psi_{u,b}\Big(\sum_{n\geqslant 0} a_n X^n\Big)\psi_{u,b}\Big(\sum_{n\geqslant 0} b_n X^n\Big).$$

Soit maintenant un autre morphisme d'anneaux $\phi : A[X] \longrightarrow B$ satisfaisant

$$\phi(a) = u(a), \quad \text{pour tout } a \in A, \quad \text{et} \quad \phi(X) = b.$$

On a alors (puisque ϕ est un morphisme d'anneaux)

$$\phi\Big(\sum_{n\geqslant 0} a_n X^n\Big) = \sum_{n\geqslant 0} \phi(a_n)\big(\phi(X)\big)^n = \sum_{n\geqslant 0} u(a_n)b^n = \psi_{u,b}\Big(\sum_{n\geqslant 0} a_n X^n\Big),$$

et donc $\phi = \psi_{u,b}$.

Enfin, soit $\varphi : A[X] \longrightarrow B$ un morphisme d'anneaux. Si $\varphi = \psi_{u,b}$, par définition de $\psi_{u,b}$, on a nécessairement $u = \varphi_{|A}$ et $b = \varphi(X)$. Ainsi, si u et b existent, ils sont uniques.

Posons $u = \varphi_{|A}$ et $b = \varphi(X)$, et montrons que $\varphi = \psi_{u,b}$. Puisque φ est un morphisme d'anneaux, il en est de même de $u : A \longrightarrow B$. On vient de voir que l'on a

$$\varphi\Big(\sum_{n\geqslant 0} a_n X^n\Big) = \sum_{n\geqslant 0} \varphi(a_n)\big(\varphi(X)\big)^n.$$

Par définition de u et b, cela s'écrit

$$\varphi\Big(\sum_{n\geqslant 0} a_n X^n\Big) = \sum_{n\geqslant 0} u(a_n)b^n = \psi_{u,b}\Big(\sum_{n\geqslant 0} a_n X^n\Big).$$

Ainsi, $\varphi = \psi_{u,b}$. □

6.17. Exemple. Soit B un anneau, que l'on ne suppose pas commutatif. Soit A un sous-anneau de B et $b \in B$. Pour tout $f = \sum_{n\geqslant 0} a_n X^n \in A[X]$, on pose

$$f(b) = \sum_{n\geqslant 0} a_n b^n.$$

L'application

$$\begin{aligned} A[X] &\longrightarrow B \\ f = \sum_{n \geqslant 0} a_n X^n &\longmapsto f(b) \end{aligned}$$

est appelée *évaluation en b*. On montre aisément que c'est un morphisme de groupes. En revanche, ce n'est pas en général un morphisme d'anneaux (cf. remarque 6.22).

C'est en revanche le cas si l'anneau B est commutatif, comme on le voit en appliquant la proposition 6.16 à l'inclusion $A \subset B$.

6.18. Remarque. Si $f \in A[X]$, on peut en particulier définir une fonction polynomiale

$$\begin{aligned} \widetilde{f} \colon A &\longrightarrow A \\ a &\longmapsto f(a), \end{aligned}$$

mais cette fonction ne caractérise pas nécessairement le polynôme f. Plus précisément, l'application

$$\begin{aligned} A[X] &\longrightarrow \mathscr{F}(A, A) \\ f &\longmapsto \widetilde{f} \end{aligned}$$

n'est pas nécessairement injective. Par exemple, si $A = \mathbb{F}_2$, la fonction polynomiale associée au polynôme $f = X^2 - X$ est identiquement nulle.

L'exemple 6.17 nous conduit tout naturellement à nous intéresser au noyau du morphisme de spécialisation. Cela conduit à la notion de racine d'un polynôme, notion extrêmement importante qui sera étudiée tout particulièrement dans la partie sur la théorie des corps.

6.19. Définition. Soit A un anneau (non nécessairement commutatif), et soit $f \in A[X]$.

On dit que $a \in A$ est une *racine* (ou un *zéro*) de f si $f(a) = 0$.

6.20. Lemme. *Soit A un anneau commutatif. Soit $f \in A[X]$ un polynôme et soit $a \in A$. Alors, a est une racine de f si, et seulement si, il existe un polynôme $g \in A[X]$ tel que $f = (X - a)g$.*

Démonstration. D'après la proposition 6.16, il existe $Q, R \in A[X]$ tels que

$$f = (X - a)Q + R,$$

où $\deg(R) < 1$. Autrement dit, R est un polynôme constant, et l'on a donc

$$f = (X - a)Q + r, r \in A.$$

Puisque A est commutatif, l'évaluation en a

$$\begin{aligned} A[X] &\longrightarrow A \\ P &\longmapsto P(a) \end{aligned}$$

est un morphisme d'anneaux. On en déduit que $f(a) = (a-a)Q(a)+r = r$. Par conséquent, $f(a) = 0$ si, et seulement si, $r = 0$, ce qui entraîne le résultat voulu. □

Il est bien connu que si K est un corps, et si $f \in K[X]$ est un polynôme non nul, alors f possède au plus $\deg(f)$ racines dans K. Cela se généralise aisément à un anneau intègre (voir le lemme ci-dessous), mais cela devient faux en général, puisque $\overline{2}X \in \mathbb{Z}/4\mathbb{Z}[X]$ admet deux racines.

En fait, on a le résultat suivant.

6.21. Lemme. *Soit A un anneau commutatif non trivial. Alors, tout polynôme non nul $f \in A[X]$ possède au plus $\deg(f)$ racines dans A si, et seulement si, A est intègre.*

Démonstration. Supposons que A soit intègre. Nous allons démontrer le résultat par récurrence sur le degré d de f. Si $d = 0$, f est un polynôme constant non nul, et n'a aucune racine dans A. Supposons avoir démontré que tout polynôme non nul de degré $d \geqslant 0$ possède au plus d racines dans A et soit f un polynôme non nul de degré $d+1$. Si f n'a pas de racines dans A, on a fini. Supposons donc que f possède au moins une racine $a \in A$. Par le lemme 6.20, il existe $g \in A[X]$ tel que $f = (X-a)g$. Puisque A est intègre, on a

$$\deg\big((X-a)g\big) = \deg(X-a) + \deg(g) = 1 + \deg(g).$$

Ainsi, $\deg(g) = d$ et par hypothèse de récurrence g a au plus d racines. Pour conclure, il suffit de constater qu'une racine de f est soit égale à a, soit une racine de g.

En effet, si $b \in A$ vérifie $f(b) = 0$, alors l'évaluation en b (qui est un morphisme d'anneaux, car A est commutatif) nous donne

$$0 = (b-a)g(b).$$

Par intégrité de A, on en déduit que $b = a$ ou $g(b) = 0$, ce qu'il fallait vérifier. On en déduit donc que f a au plus $d+1$ racines, ce qui achève la récurrence.

Supposons maintenant que A ne soit pas intègre. Puisque A est commutatif et non trivial, cela implique l'existence de deux éléments $a, b \in A$ non nuls tels que $ab = 0$. Mais alors, le polynôme aX possède au moins deux racines distinctes (qui sont 0 et b). □

6.22. Remarque. La commutativité de A est cruciale. En effet, si A n'est pas commutatif, les résultats précédents deviennent faux, même si A n'a pas de diviseurs de zéro. Par exemple, soit $\mathbb{H} \subset \mathrm{M}_2(\mathbb{C})$ l'ensemble défini par

$$\mathbb{H} = \left\{ M \in \mathrm{M}_2(\mathbb{C}) \;\middle|\; M = \begin{pmatrix} z_1 & -z_2 \\ \overline{z}_2 & \overline{z}_1 \end{pmatrix},\ z_1, z_2 \in \mathbb{C} \right\}.$$

On vérifie sans mal que $\mathbb{H}$ est un anneau à division (cf. exercice 20). En particulier, $\mathbb{H}$ est sans diviseurs de zéro.

Posons

$$e = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad f = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Alors, le polynôme $X^2 + \mathrm{I}_2 \in \mathbb{H}[X]$ possède au moins 4 racines dans $\mathbb{H}$, qui sont $e, -e, f, -f$ (en fait, il en admet même une infinité). Cela s'explique par le fait que la non-commutativité de $\mathbb{H}$ implique que l'évaluation d'un polynôme en un élément de $\mathbb{H}$ n'est plus un morphisme d'anneaux (et donc la démonstration précédente n'est plus valable).

Par exemple, on a $X^2 + \mathrm{I}_2 = (X - e)(X + e), f^2 + \mathrm{I}_2 = 0$, mais on vérifie sans peine que $(f - e)(f + e) \neq 0$, car $fe \neq ef$.

En particulier, la relation entre factorisation et racines n'a plus lieu. Par exemple, on a $X^2 + \mathrm{I}_2 \neq (X - e)(X - f)$.

Nous allons maintenant continuer notre étude des zéros des polynômes. Lorsque f est une fonction polynomiale à valeurs dans $\mathbb{R}$, les dérivées successives de f permettent d'en détecter les racines multiples. Dans le cas des polynômes abstraits, on peut également définir une notion de dérivée formelle bien pratique, qui jouit peu ou prou des mêmes propriétés.

6.23. Définition. Soit A un anneau commutatif.

Le *polynôme dérivé* de $P = \sum_{n \geqslant 0} a_n X^n \in A[X]$, noté P', est le polynôme

$$P' = \sum_{n \geqslant 1} n \cdot a_n X^{n-1}.$$

On a alors le résultat suivant, dont la démonstration est laissée au lecteur à titre d'exercice.

6.24. Lemme. *Soit A un anneau commutatif.*

(1) *Si P est constant, alors $P' = 0$. Si de plus A est de caractéristique nulle, pour tout $P \in A[X]$, on a $P' = 0$ si, et seulement si, P est constant ;*

(2) *pour tous $P, Q \in A[X]$, on a $(P+Q)' = P'+Q'$ et $(PQ)' = P'Q+PQ'$.*

6.25. Remarque. La réciproque du (1) est fausse sans hypothèse sur A. En effet, la dérivée de $X^2 \in \mathbb{F}_2[X]$ est nulle, bien que X^2 soit non constant.

Notation. Si $P \in A[X]$, et $k \geqslant 0$, on notera $P^{(k)}$ le polynôme obtenu en dérivant k fois formellement. Par convention, $P^{(0)} = P$.

On définit maintenant la notion d'ordre d'une racine d'un polynôme.

6.26. Définition. Soit A un anneau commutatif, soit $P \in A[X]$. On dit que $a \in A$ est une racine de P *d'ordre* s s'il existe $Q \in A[X]$ tel que :

(1) on a $P = (X-a)^s Q$;

(2) on a $Q(a) \neq 0$.

Autrement dit, a est une racine d'ordre s si, et seulement si, $(X-a)^s$ est la plus grande puissance de $X-a$ divisant [2] P.

Si $P = 0$, on convient que tout élément $a \in A$ est une racine de P, d'ordre infini.

On dit qu'une racine de P est *simple* si elle est d'ordre 1, et *multiple* sinon.

6.27. Remarque. Avec cette définition, une « racine d'ordre 0 » est un élément $a \in A$ tel que $P(a) \neq 0$. Ce n'est donc pas une racine de P. Inclure le cas $s = 0$ est néanmoins bien pratique dans les énoncés, et évite de devoir considérer des cas particuliers, comme par exemple dans le résultat suivant.

6.28. Lemme. *Soit K un corps de caractéristique nulle. Soit $P \in K[X]$, et soit $a \in K$ une racine de P d'ordre $s \geqslant 1$. Alors, a est une racine de P' d'ordre $s-1$.*

Démonstration. Écrivons $P = (X-a)^s Q$, avec $Q(a) \neq 0$. Alors,

$$P' = (X-a)^{s-1}(sQ + (X-a)Q').$$

Si l'on avait $(sQ+(X-a)Q')(a) = 0$, on obtiendrait $sQ(a) = 0$. Comme K est un corps, $Q(a)$ est inversible, et l'on a alors $s \cdot 1_K = 0_K$, d'où une contradiction, puisque K est de caractéristique nulle. Ainsi, a est une racine de P' de multiplicité $s-1$. □

On a aussi une caractérisation de l'ordre d'une racine en terme de dérivées successives.

2. Dans un sens évident.

6.29. Lemme. *Soit K un corps de caractéristique nulle. Alors, $a \in K$ est une racine de $P \in K[X]$ d'ordre $s \geqslant 0$ si, et seulement si :*

(1) *pour tout $k \in [\![1, s-1]\!]$, on a $P^{(k)}(a) = 0$;*

(2) *on a $P^{(s)}(a) \neq 0$.*

Démonstration. En écrivant $X = (X - a) + a$ et en utilisant la formule du binôme de Newton (valable dans tout anneau commutatif), on voit que tout $P \in K[X]$ peut s'écrire sous la forme

$$P = c_n(X-a)^n + \cdots + c_1(X-a) + c_0, \quad c_i \in K.$$

Un simple calcul montre alors que, pour tout $k \in [\![0, n]\!]$, on a

$$k!\, c_k = P^{(k)}(a).$$

Puisque K est un corps de caractéristique nulle, $(k!)\cdot 1_K$ est non nul, donc inversible dans K, et l'on a $c_k = \dfrac{P^{(k)}(a)}{k!}$. On obtient ainsi l'égalité [3]

$$P = \frac{P^{(n)}(a)}{n!}(X-a)^n + \cdots + \frac{P'(a)}{1!}(X-a) + P(a).$$

Cela démontre en particulier l'unicité de l'écriture précédente.

Ainsi, on peut écrire

$$P = (X-a)^s Q + \frac{P^{(s-1)}(a)}{(s-1)!}(X-a)^{s-1} + \cdots + \frac{P'(a)}{1!}(X-a) + P(a),$$

avec $Q \in K[X]$.

Autrement dit, le polynôme

$$R = \frac{P^{(s-1)}(a)}{(s-1)!}(X-a)^{s-1} + \cdots + \frac{P'(a)}{1!}(X-a) + P(a)$$

est le reste de la division euclidienne de P par $(X-a)^s$.

Par unicité de l'écriture selon les puissances de $X - a$, on obtient que le polynôme $(X-a)^s$ divise P si, et seulement si, on a

$$P(a) = P'(a) = \cdots = P^{(s-1)}(a) = 0.$$

On en déduit aisément que $(X-a)^s$ est la plus grande puissance de $X - a$ divisant P si, et seulement si, les conditions du lemme sont vérifiées. □

3. Ici, on a fait l'abus de notation bien pratique $\dfrac{1}{m!} = (m!\, 1_K)^{-1}$.

6.30. Remarque. Sans l'hypothèse sur K, les résultats précédents deviennent faux. Par exemple, si $K = \mathbb{F}_2$, alors 0 est une racine d'ordre 2 du polynôme $P = X^2$, mais toutes les dérivées successives de P sont nulles.

Nous allons maintenant nous intéresser aux diviseurs de zéro et aux éléments inversibles de $A[X]$. Commençons par un lemme facile.

6.31. Lemme. *Soit A un anneau intègre. Alors, on a les propriétés suivantes :*

(1) *l'anneau $A[X]$ n'a pas de diviseurs de zéro ;*

(2) *on a $A[X]^\times = A^\times$;*

(3) *on a $\mathrm{Nil}(A[X]) = (0)$.*

Démonstration. Puisque $A[X]$ est intègre par le lemme 6.9, il n'a pas de diviseurs de zéro, donc en particulier pas d'éléments nilpotents non nuls d'après la remarque 2.9, d'où les points (1) et (3). Il reste à montrer le point (2). Si $a \in A^\times$, alors $a \in A[X]^\times$, puisque les relations

$$aa^{-1} = a^{-1}a = 1 \in A$$

sont également vraies dans $A[X]$. Réciproquement, soit $f \in A[X]^\times$. Alors, il existe $g \in A[X]$ tel que $fg = gf = 1$. En particulier, f et g sont non nuls. Comme A est intègre, on a

$$\deg(fg) = \deg(f) + \deg(g) = \deg(1) = 0$$

par le lemme 6.9. Mais le degré d'un polynôme non nul étant un entier positif ou nul, on en déduit

$$\deg(f) = \deg(g) = 0,$$

donc $f = a$ et $g = b$ pour certains $a, b \in A$. Mais alors, on a $ab = ba = 1$, et donc $f = a \in A^\times$. □

Ce résultat devient faux lorsque A est quelconque.

6.32. Exemple. On a $\overline{2}X \neq 0 \in \mathbb{Z}/4\mathbb{Z}[X]$, mais $(\overline{2}X)(\overline{2}X) = 0$. De plus, on vérifie que $(\overline{1} + \overline{2}X)^2 = \overline{1}$ et donc $\overline{1} + \overline{2}X$ est inversible dans $\mathbb{Z}/4\mathbb{Z}[X]$, mais non constant [4].

Le but de ce qui suit est de calculer les éléments inversibles, les éléments nilpotents et les diviseurs de zéro de $A[X]$.

4. On peut aussi appliquer le lemme 2.11 dans l'anneau $\mathbb{Z}/4\mathbb{Z}[X]$, en remarquant que $\overline{2}X$ est nilpotent, puisque $(\overline{2}X)^2 = 0$.

On commence par déterminer les diviseurs de zéro.

6.33. Proposition. *Soit A un anneau commutatif.*

Alors, un polynôme $f \in A[X]$ non nul est diviseur de zéro si, et seulement si, il existe $d \in A \setminus \{0\}$ tel que $df = 0$.

Autrement dit, si $f = a_n X^n + \cdots + a_1 X + a_0$, $a_n \in A \setminus \{0\}$, f est diviseur de zéro si, et seulement si, il existe $d \in A \setminus \{0\}$ tel que

$$da_i = 0, \quad \text{pour tout } i \in [\![0, n]\!].$$

Démonstration. S'il existe $d \in A \setminus \{0\}$ tel que $df = 0$, alors f est bien sûr diviseur de zéro.

Inversement, supposons que $f = a_n X^n + \cdots + a_1 X + a_0$, $a_m \in A \setminus \{0\}$, soit diviseur de zéro. Alors, il existe $g \in A[X]$ non nul tel que $fg = 0$. Choisissons le polynôme g de sorte que son degré $m \geqslant 0$ soit minimal, et écrivons

$$g = b_m X^m + \cdots + b_1 X + b_0, \quad b_m \in A \setminus \{0\}.$$

Montrons par l'absurde que $m = 0$. Supposons au contraire que $m \geqslant 1$. En particulier, on a $df \neq 0$ pour tout $d \in A \setminus \{0\}$.

Si $a_i g = 0$ pour tout $i \in [\![0, n]\!]$, alors, en comparant les coefficients dominants, on obtient $a_i b_m = 0$ pour tout $i \in [\![0, n]\!]$. Mais alors, $b_m f = 0$, d'où une contradiction. Ainsi, il existe $i \in [\![0, n]\!]$ tel que $a_i g \neq 0$. Soit $\ell \in [\![0, n]\!]$ le plus grand entier tel que $a_\ell g \neq 0$. Alors, on a

$$0 = fg = (a_\ell X^\ell + \cdots + a_0)(b_m X^m + \cdots + b_0).$$

Par conséquent, $a_\ell b_m = 0$ et donc

$$a_\ell g = a_\ell b_{m-1} X^{m-1} + \cdots + a_\ell b_0.$$

Ainsi, $\deg(a_\ell g) < m$. De plus, on a $a_\ell g \neq 0$ par choix de ℓ. Mais on a

$$f(a_\ell g) = a_\ell f g = 0,$$

ce qui contredit la minimalité de m. Cela achève la démonstration. □

6.34. Remarque. L'ensemble des diviseurs de zéro n'est pas égal à l'ensemble des polynômes dont les coefficients sont soit nuls, soit diviseurs de zéro. Par exemple, si $A = \mathbb{Z}/6\mathbb{Z}$, le polynôme $\overline{2}X + \overline{3} \in A[X]$ n'est pas diviseur de zéro au vu de la proposition précédente, bien que $\overline{2}$ et $\overline{3}$ soient des diviseurs de zéro.

Nous allons maintenant décrire les éléments nilpotents de $A[X]$.

6.35. Proposition. *Soit A un anneau commutatif.*

Alors, le polynôme $a_nX^n + \cdots + a_1X + a_0 \in A[X]$ est nilpotent si, et seulement si, $a_0, \ldots, a_n \in \mathrm{Nil}(A)$. Autrement dit, on a

$$\mathrm{Nil}(A[X]) = \mathrm{Nil}(A)[X],$$

avec des notations évidentes.

Démonstration. Soit $f = a_nX^n + \cdots + a_1X + a_0 \in A[X]$.

Supposons tout d'abord que $a_0, \ldots, a_n \in \mathrm{Nil}(A)$. Il découle de la définition que $\mathrm{Nil}(A) \subset \mathrm{Nil}(A[X])$ Puisque $\mathrm{Nil}(A[X])$ est un idéal de $A[X]$ par le lemme 4.16 (1), on a alors $f \in \mathrm{Nil}(A[X])$.

Inversement, supposons que $f \in \mathrm{Nil}(A[X])$. Soit $r \geqslant 1$ tel que $f^r = 0$. En comparant les coefficients dominants, on obtient $a_n^r = 0$. Autrement dit, $a_n \in \mathrm{Nil}(A)$, et par suite $f - a_nX^n \in \mathrm{Nil}(A[X])$, car $\mathrm{Nil}(A[X])$ est un idéal. Or, $f - a_nX^n = a_{n-1}X^{n-1} + \cdots + a_1X + a_0$. En appliquant le raisonnement précédent à ce nouveau polynôme, on obtient $a_{n-1} \in \mathrm{Nil}(A)$. Par récurrence descendante, on obtient finalement que $a_0, \ldots, a_n \in \mathrm{Nil}(A)$. Cela achève la démonstration. □

Continuons en déterminant les idempotents de $A[X]$.

6.36. Proposition. *Soit A un anneau commutatif. Alors, les idempotents de $A[X]$ sont exactement les idempotents de A.*

Démonstration. Clairement, un idempotent de A est aussi un idempotent de $A[X]$. Inversement, soit $f = a_nX^n + \cdots + a_1X + a_0 \in A[X]$ un polynôme tel que $f^2 = f$. En comparant les termes constants, on obtient $a_0^2 = a_0$.

Montrons par récurrence finie que $a_k = 0$ pour tout $k \in [\![1, n]\!]$. Pour tout entier $k \in [\![1, n]\!]$, soit (H_k) la propriété : (H_k) $a_i = 0$ pour tout $i \in [\![1, k]\!]$.

En comparant les termes en X dans l'égalité $f^2 = f$, on obtient $2a_0a_1 = a_1$. On a donc $2a_0^2a_1 = 2a_0a_1 = a_0a_1$, soit $a_0a_1 = 0$. Mais alors, $a_1 = 0$, et par conséquent, (H_1) est vraie. Supposons que la propriété (H_k) soit vraie pour un entier $k \in [\![1, n-1]\!]$. On a donc $f = a_nX^n + \cdots + a_{k+1}X^{k+1} + a_0$. En comparant cette fois les termes en X^{k+1}, on obtient $2a_0a_{k+1} = a_{k+1}$, et par un raisonnement similaire au précédent, on a $a_{k+1} = 0$. Ainsi, (H_{k+1}) est vraie.

Par récurrence finie, on obtient que (H_n) est vraie. Par conséquent, $f = a_0$. Comme a_0 est un idempotent de A, cela achève la démonstration. □

On finit ce chapitre en déterminant les unités de $A[X]$ dans le cas général.

6.37. Proposition. *Soit A un anneau commutatif.*

Alors, le polynôme $a_nX^n + \cdots + a_1X + a_0 \in A[X]$ est inversible si, et seulement si, $a_0 \in A^\times$ et $a_1, \ldots, a_n \in \mathrm{Nil}(A)$. Autrement dit, on a

$$A[X]^\times = A^\times + X\mathrm{Nil}(A)[X].$$

Démonstration. Soit $f = a_nX^n + \cdots + a_1X + a_0$. On suppose que $a_0 \in A^\times$ et que $a_1, \ldots, a_n \in \mathrm{Nil}(A)$. Comme $\mathrm{Nil}(A[X])$ est un idéal de $A[X]$, on a

$$a_nX^n + \cdots + a_1X \in \mathrm{Nil}(A[X]).$$

Par le lemme 4.16 (2), on a $f \in A[X]^\times$ puisque $a_0 \in A^\times \subset A[X]^\times$.

Inversement, supposons que $f = a_nX^n + \cdots + a_1X + a_0 \in A[X]^\times$.

Si f est constant, alors $f \in A^\times$ et le résultat est clair. Supposons f non constant, et soit $g = b_mX^m + \cdots + b_1X + b_0 \in A[X]$ tel que $fg = 1$.

On a en particulier $f(0)g(0) = a_0b_0 = 1$, et par suite a_0, $b_0 \in A^\times$. Puisque f est non constant, on a $n + m \geqslant 1$. En comparant les termes en X^{n+m}, on obtient alors $a_nb_m = 0$.

Montrons par récurrence sur i que $a_n^i b_{m+1-i} = 0$, pour tout $i \in [\![1, m+1]\!]$. C'est vrai pour $i = 1$ d'après le point précédent. Supposons avoir démontré les égalités

$$a_nb_m = 0, \quad a_n^2 b_{m-1} = 0, \quad \ldots, \quad a_n^i b_{m+1-i} = 0.$$

En multipliant chaque égalité par une puissance a_n convenable, on obtient alors

$$a_n^i b_m = a_n^i b_{m-1} = \cdots = a_n^i b_{m+1-i} = 0.$$

En regardant les termes en X^{n+m-i} dans l'égalité $fg = 0$, on obtient

$$a_nb_{m-i} + a_{n-1}b_{m-i+1} + \cdots + a_{n-i}b_m = 0.$$

En multipliant par a_n^i et en utilisant les égalités ci-dessus, on en déduit l'égalité

$$a_n^{i+1} b_{m-i} = 0,$$

ce qui achève la récurrence. En particulier, $a_n^{m+1} b_0 = 0$, et comme $b_0 \in A^\times$, on obtient $a_n^{m+1} = 0$.

Ainsi, $a_n \in \mathrm{Nil}(A)$, et par suite $a_nX^n \in \mathrm{Nil}(A[X])$, car $\mathrm{Nil}(A[X])$ est un idéal. Par le lemme 4.16 (2), on a

$$f - a_nX^n = a_{n-1}X^{n-1} + \cdots + a_1X + a_0 \in A[X]^\times.$$

En appliquant le raisonnement précédent à ce dernier polynôme, on obtient alors $a_{n-1} \in \mathrm{Nil}(A)$. Par récurrence descendante, on en déduit finalement que $a_1, \ldots, a_n \in \mathrm{Nil}(A)$. Cela achève la démonstration. □

6.38. Corollaire. *Soit A un anneau commutatif. Alors, $A[X]^\times = A^\times$ si, et seulement si, A n'a pas d'éléments nilpotents non nuls.*

Nous terminons ce chapitre en étudiant brièvement les polynômes symétriques en plusieurs indéterminées.

7. Polynômes symétriques

Nous allons maintenant étudier les polynômes symétriques à coefficients dans un anneau, ce qui nous permettra ensuite d'établir certains résultats classiques sur les polynômes en une variable à coefficients dans un corps dans les chapitres ultérieurs.

Soit A un anneau commutatif, et soit $P \in A[X_1, \ldots, X_n]$. Un tel polynôme P s'écrit donc

$$P = \sum_{m_1,\ldots,m_n} a_{m_1,m_2,\ldots,m_n} X_1^{m_1} \cdots X_n^{m_n},$$

où les coefficients $a_{m_1,m_2,\ldots,m_n} \in A$ sont presque tous nuls.

Nous allons maintenant introduire une notation plus commode.

Notation. Si $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$, on pose

$$X^\alpha = X_1^{\alpha_1} \cdots X_n^{\alpha_n}.$$

Un polynôme $P \in A[X_1, \ldots, X_n]$ s'écrit donc de manière unique

$$P = \sum_\alpha a_\alpha X^\alpha, \quad a_\alpha \in A,$$

où les coefficients a_α sont presque tous nuls.

Si $\tau \in \mathfrak{S}_n$, on pose

$$\tau \cdot P = P(X_{\tau(1)}, \ldots, X_{\tau(n)}).$$

Pour tout $\alpha \in \mathbb{N}^n$, on a

$$\tau \cdot X^\alpha = X_{\tau(1)}^{\alpha_1} \cdots X_{\tau(n)}^{\alpha_n} = X_1^{\alpha_{\tau^{-1}(1)}} \cdots X_n^{\alpha_{\tau^{-1}(n)}} = X^{\tau\cdot\alpha},$$

où l'action de $\mathfrak{S}_n$ sur $\mathbb{N}^n$ est donnée par

$$\begin{aligned} \mathfrak{S}_n \times \mathbb{N}^n &\longrightarrow \mathbb{N}^n \\ \big(\tau, (\alpha_1, \ldots, \alpha_n)\big) &\longmapsto (\alpha_{\tau^{-1}(1)}, \ldots, \alpha_{\tau^{-1}(n)}). \end{aligned}$$

Si $P = \sum_\alpha a_\alpha X^\alpha$, on a donc

$$\tau \cdot P = \sum_\alpha a_\alpha X^{\tau\cdot\alpha} = \sum_\alpha a_{\tau^{-1}\cdot\alpha} X^\alpha.$$

On en déduit facilement que l'application

$$\begin{aligned} \mathfrak{S}_n \times A[X_1, \ldots, X_n] &\longrightarrow A[X_1, \ldots, X_n] \\ (\tau, P) &\longmapsto \tau \cdot P \end{aligned}$$

définit une action de $\mathfrak{S}_n$ sur $A[X_1, \ldots, X_n]$ telle que

$$\tau \cdot (P_1 + aP_2) = \tau \cdot P_1 + a(\tau \cdot P_2) \quad \text{et} \quad \tau \cdot (P_1 P_2) = (\tau \cdot P_1)(\tau \cdot P_2),$$

pour tous P_1, $P_2 \in A[X_1, \ldots, X_n]$, tout $a \in A$, et tout $\tau \in \mathfrak{S}_n$.

7.1. Définition. Soit A un anneau commutatif, et soit $P \in A[X_1, \ldots, X_n]$. On dit que P est *symétrique* s'il est fixe sous l'action de $\mathfrak{S}_n$ décrite précédemment, i.e. si

$$P(X_{\tau(1)}, \ldots, X_{\tau(n)}) = P(X_1, \ldots, X_n), \quad \text{pour tout } \tau \in \mathfrak{S}_n.$$

L'ensemble des polynômes symétriques est un sous-anneau de $A[X_1, \ldots, X_n]$, comme on le vérifie aisément.

7.2. Exemples

(1) Si $n = 3$, $X_1 + X_2 + X_3$ est symétrique.

(2) Si $n = 4$, $X_1 + X_2 + X_3$ n'est **PAS** symétrique, comme on le voit en application la transposition (1 4), par exemple.

(3) Si $n = 3$, $X_1 X_2 X_3$ est symétrique.

Le lemme suivant donne des exemples importants de polynômes symétriques.

7.3. Lemme. *Soit A un anneau commutatif et soit $n \geqslant 1$. Alors, pour tout entier $j \in [\![1, n]\!]$, le polynôme*

$$\sigma_j = \sum_{1 \leqslant i_1 < i_2 < \cdots < i_j \leqslant n} X_{i_1} X_{i_2} \cdots X_{i_j} \in A[X_1, \ldots, X_n]$$

est symétrique.

Démonstration. Pour tout $j \in [\![1, n]\!]$, on note $\mathscr{P}_j(n)$ l'ensemble des parties de $[\![1, n]\!]$ à j éléments. Si $I \in \mathscr{P}_j(n)$, on pose

$$X_I = \prod_{i \in I} X_i.$$

Avec ces notations, on a donc

$$\sigma_j = \sum_{I \in \mathscr{P}_j(n)} X_I.$$

Remarquons maintenant que pour tout $I \in \mathscr{P}_j(n)$, et tout $\tau \in \mathfrak{S}_n$, on a

$$\tau \cdot X_I = \prod_{i \in I} X_{\tau(i)} = \prod_{k \in \tau(I)} X_k = X_{\tau(I)}.$$

Comme τ est bijective, l'application

$$\begin{aligned} \mathscr{P}_j(n) &\longrightarrow \mathscr{P}_j(n) \\ I &\longmapsto \tau(I) \end{aligned}$$

est bijective. Ainsi, on obtient

$$\tau \cdot \sigma_j = \sum_{I \in \mathscr{P}_j(n)} X_{\tau(I)} = \sum_{I \in \mathscr{P}_j(n)} X_I = \sigma_j,$$

d'où le résultat. □

7.4. Définition. Soit $j \in [\![1, n]\!]$. Le polynôme

$$\sigma_j = \sum_{1 \leqslant i_1 < i_2 < \cdots < i_j \leqslant n} X_{i_1} X_{i_2} \cdots X_{i_j} \in A[X_1, \ldots, X_n]$$

est appelé le j-ème *polynôme symétrique élémentaire de* $A[X_1, \ldots, X_n]$.

7.5. Exemple. Pour tout $n \geqslant 1$, on a par exemple

$$\sigma_1 = X_1 + X_2 + \cdots + X_n \quad \text{et} \quad \sigma_n = X_1 X_2 \cdots X_n.$$

Clairement, toute expression polynomiale en les polynômes symétriques élémentaires est un polynôme symétrique. On va maintenant voir que la réciproque est vraie. Auparavant, on a besoin de quelques définitions.

7.6. Définition. Un élément de $A[X_1, \cdots, X_n]$ de la forme

$$aX^\alpha = aX_1^{\alpha_1} \cdots X_n^{\alpha_n}, \quad a \in A, \ \alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$$

est appelé un *monôme.*

Le *degré* d'un monôme $aX_1^{m_1} \cdots X_n^{m_n}$, $a \in A \setminus \{0\}$ est $m_1 + \cdots + m_n$. Le *degré* d'un polynôme $P \in A[X_1, \ldots, X_n]$ non nul est le maximum des degrés des monômes non nuls qui le constituent.

On dit que $P \in A[X_1, \cdots, X_n]$ non nul est *homogène de degré* $d \geqslant 0$ s'il est somme de monômes de même degré d. Par convention [5], le polynôme nul est homogène de degré d pour tout $d \geqslant 0$.

Clairement, tout polynôme $P \in A[X_1, \ldots, X_n]$ non nul peut être écrit de manière unique comme somme de polynômes homogènes non nuls de degré deux à deux distincts. On laisse le lecteur consciencieux vérifier ce fait à titre d'exercice.

5. Ou parce qu'une somme vide est nulle.

7.7. Exemple. Pour tout $j \in [\![1, n]\!]$, le polynôme $\sigma_j \in A[X_1, \ldots, X_n]$ est homogène de degré j.

Avant d'entrer dans le vif du sujet, nous allons définir une autre notion de degré pour les polynômes symétriques.

Notation. Si $P \in A[X_1, \ldots, X_n]$, on note $\deg_{X_i}(P)$ le degré de P vu comme un élément de $A[X_1, \ldots, X_{i-1}, X_{i+1}, \ldots, X_n][X_i]$.

On a alors le lemme suivant.

7.8. Lemme. *Soit A un anneau commutatif, et soit un polynôme symétrique $P \in A[X_1, \ldots, X_n]$. Alors, $\deg_{X_i}(P)$ est indépendant de l'entier i.*

Démonstration. Soient $i, j \in [\![1, n]\!]$, avec $i \neq j$, et soit $\tau = (i\ j) \in \mathfrak{S}_n$. Posons $m_i = \deg_{X_i}(P)$ et $m_j = \deg_{X_j}(P)$. Par hypothèse, P contient un monôme dont la puissance en X_i est égale à m_i. Mais alors, $P = \tau \cdot P$ contient un monôme dont la puissance en X_j est égale à m_i. Par définition de m_j, on en déduit $m_i \leqslant m_j$. En échangeant les rôles de i et j, on obtient également $m_j \leqslant m_i$. Ainsi, $m_i = m_j$, d'où le résultat voulu. □

7.9. Définition. Soit A un anneau commutatif, et soit un polynôme symétrique $P \in A[X_1, \ldots, X_n]$. Le *degré partiel* de P est l'entier $\deg_{X_i}(P)$, où i est un élément arbitraire de $[\![1, n]\!]$. On le note $\omega(P)$.

7.10. Exemple. Chaque polynôme symétrique élémentaire est de degré partiel égal à 1.

7.11. Théorème (Structure des polynômes symétriques). *Soit A un anneau commutatif. Pour tout polynôme symétrique $P \in A[X_1, \ldots, X_n]$, il existe un unique polynôme $Q \in A[T_1, \ldots, T_n]$ tel que*

$$P = Q(\sigma_1, \ldots, \sigma_n).$$

De plus, si P est de degré partiel égal à d, alors Q est de degré d.

On va avoir besoin de plusieurs résultats intermédiaires.

7.12. Proposition. *Soit A un anneau commutatif et soit $n \geqslant 1$. Alors, pour tout $Q \in A[T_1, \ldots, T_n]$, on a*

$$Q(\sigma_1, \ldots, \sigma_n) = 0 \quad \Longrightarrow \quad Q = 0.$$

Démonstration

On procède par récurrence sur le nombre n d'indéterminées.

Si $n = 1$, c'est clair, puisque $\sigma_1(X_1) = X_1$.

On suppose que la propriété voulue est vraie pour les polynômes S en n variables, et l'on suppose qu'elle est **fausse** pour les polynômes en $n+1$ variables. On choisit alors un polynôme $S \neq 0$ de degré minimal tel que

$$S(\sigma_1, \cdots, \sigma_{n+1}) = 0. \qquad (*)$$

où $\sigma_1, \ldots, \sigma_{n+1}$ sont les polynômes symétriques élémentaires en les indéterminées $X_1, \ldots, X_{n+1}$. Exceptionnellement, on note $\Sigma_1, \cdots, \Sigma_n$ les polynômes symétriques élémentaires en $X_1, \ldots, X_n$. Remarquons que l'on a

$$\sigma_j(X_1, \ldots, X_n, 0) = \Sigma_j, \quad \text{pour tout } j \in [\![1, n]\!], \qquad (**)$$

$$\sigma_{n+1}(X_1, \ldots, X_n, 0) = 0. \qquad (***)$$

Écrivons $S = S_m T_{n+1}^m + \cdots + S_1 T_{n+1} + S_0$, $S_0, \ldots, S_m \in A[T_1, \cdots, T_n]$.

Supposons tout d'abord que $S_0 = 0$. On peut alors écrire $S = T_{n+1}U, U \neq 0$. Il est alors clair que U est de degré strictement inférieur à celui de S. Par hypothèse sur S, on a

$$\sigma_{n+1} U(\sigma_1, \cdots, \sigma_{n+1}) = 0.$$

Or, la multiplication par X_i dans $A[X_1, \ldots, X_{n+1}]$ est injective. Comme on a $\sigma_{n+1} = X_1 \cdots X_{n+1}$, on en déduit que $U(\sigma_1, \cdots, \sigma_{n+1}) = 0$, ce qui contredit la minimalité du degré de S.

Ainsi, $S_0 \neq 0$. D'après $(*)$, on obtient

$$S_m(\sigma_1, \ldots, \sigma_n)\sigma_{n+1}^m + \cdots + S_1(\sigma_1, \ldots, \sigma_n)\sigma_{n+1} + S_0(\sigma_1, \ldots, \sigma_n) = 0.$$

En évaluant maintenant X_{n+1} en 0 dans cette expression et en utilisant les relations $(**)$ et $(***)$, on obtient $S_0(\Sigma_1, \ldots, \Sigma_n) = 0$. Par hypothèse de récurrence, on a $S_0 = 0$, d'où une contradiction. □

En appliquant cette proposition au polynôme $Q_1 - Q_2$, on obtient la partie unicité du théorème 7.11.

7.13. Corollaire. *Soit A un anneau commutatif et soit $n \geqslant 1$. Alors, pour tous Q_1, $Q_2 \in A[T_1, \ldots, T_n]$, on a*

$$Q_1(\sigma_1, \ldots, \sigma_n) = Q_2(\sigma_1, \ldots, \sigma_n) \quad \Longrightarrow \quad Q_1 = Q_2.$$

Avant de continuer, nous allons avoir besoin de quelques définitions et notations.

Dans la suite, on ordonne $\mathbb{N}^n$ grâce à l'ordre (total) lexicographique. Autrement dit, si $\alpha = (\alpha_1, \cdots, \alpha_n)$ et $\beta = (\beta_1, \cdots, \beta_n)$ sont deux éléments de $\mathbb{N}^n$, on a $\beta \leqslant \alpha$ si $\beta_1 < \alpha_1$, ou s'il existe $k \in [\![1, n]\!]$ tel que $\beta_1 = \alpha_1, \cdots, \beta_k = \alpha_k$ et $\beta_{k+1} < \alpha_{k+1}$.

Remarquons que si $k = n$, la dernière partie de la condition est vide et les éléments α et β sont égaux.

7.14. Définition. Soit $P = \sum_{\alpha} a_\alpha X^\alpha \in A[X_1, \ldots, X_n]$.

Le *support* de P, noté $\mathrm{Supp}(P)$, est l'ensemble

$$\mathrm{Supp}(P) = \{\alpha \in \mathbb{N}^n \mid a_\alpha \neq 0\}^{6}.$$

On a bien sur l'égalité

$$P = \sum_{\alpha \in \mathrm{Supp}(P)} a_\alpha X^\alpha,$$

et chaque n-uplet $\alpha \in \mathrm{Supp}(P)$ correspond à un unique monôme dans l'écriture de P, à savoir $a_\alpha X^\alpha$.

7.15. Lemme. *Soit A un anneau commutatif, et soit $P \in A[X_1, \ldots, X_n]$ un polynôme symétrique. Alors, le support de P est stable sous l'action du groupe symétrique $\mathfrak{S}_n$ sur $\mathbb{N}^n$.*

Démonstration. On peut toujours supposer P non nul (sinon, le support de P est vide, et il n'y a rien à faire). Écrivons $P = \sum_{\alpha} a_\alpha X^\alpha$. Pour toute permutation $\tau \in \mathfrak{S}_n$, on a

$$\tau \cdot P = \sum_{\alpha} a_{\tau^{-1}\cdot\alpha} X^\alpha.$$

Lorsque P est symétrique, on obtient donc

$$a_{\tau^{-1}\cdot\alpha} = a_\alpha, \quad \text{pour tout } \tau \in \mathfrak{S}_n \text{ et tout } \alpha \in \mathbb{N}^n,$$

ou encore

$$a_{\tau\cdot\alpha} = a_\alpha, \quad \text{pour tout } \tau \in \mathfrak{S}_n \text{ et tout } \alpha \in \mathbb{N}^n.$$

En particulier, si $a_\alpha \neq 0$, on a $a_{\tau\cdot\alpha} \neq 0$ pour tout $\tau \in \mathfrak{S}_n$, d'où le résultat voulu. □

7.16. Remarque. On peut même dire un peu plus. Ce qui précède montre que l'action de $\mathfrak{S}_n$ sur $\mathbb{N}^n$ se restreint en une action sur le support de P. De plus, si α et β sont dans la même orbite, on a $a_\alpha = a_\beta$.

Si Ω est l'ensemble des orbites correspondantes, on a alors

$$P = \sum_{\omega \in \Omega} a_\omega \Big(\sum_{\alpha \in \omega} X^\alpha \Big),$$

où $a_\omega \in A \setminus \{0\}$ est égal à a_α, α étant n'importe quel représentant de ω.

6. Attention à ne pas confondre cet ensemble avec le support d'une permutation !

Notation. Si $P \neq 0$, on note $D(P)$ le plus grand élément du support de P, pour l'ordre lexicographique. Remarquons alors que l'on peut écrire

$$P = \sum_{\alpha \leqslant D(P)} a_\alpha X^\alpha.$$

Le lemme suivant résume les propriétés de $D(P)$ qui nous serviront plus loin.

7.17. Lemme. *Soit $P \in A[X_1, \ldots, X_n]$ un polynôme symétrique non nul, et soit $D(P) = (\alpha_1, \ldots, \alpha_n)$. Alors :*

(1) *on a $\alpha_1 \geqslant \alpha_2 \geqslant \cdots \geqslant \alpha_n$;*

(2) *les entiers $\alpha_1, \ldots, \alpha_n$ se calculent de manière récursive comme suit : pour tout $j \in [\![1, n]\!]$, α_j est la plus grande puissance de X_j apparaissant dans les monômes de P de la forme*

$$a_{\alpha_1,\ldots,\alpha_{j-1},\beta_j,\ldots,\beta_n} X_1^{\alpha_1} \cdots X_{j-1}^{\alpha_{j-1}} X_j^{\beta_j} \cdots X_n^{\beta_n},$$

avec $a_{\alpha_1,\ldots,\alpha_{j-1},\beta_j,\ldots,\beta_n} \in A \setminus \{0\}$;

(3) *le degré partiel de P est égal à α_1.*

Démonstration

(1) Posons $\alpha = D(P)$, et soit $a_\alpha X^\alpha$ le monôme de P correspondant. Fixons un entier $j \in [\![1, n-1]\!]$, et soit $\tau = (j\ j+1)$.

Puisque P est symétrique, le lemme 7.15 montre que le n-uplet

$$\beta = \tau \cdot \alpha = (\alpha_1, \ldots, \alpha_{j-1}, \alpha_{j+1}, \alpha_j, \alpha_{j+2}, \ldots, \alpha_n)$$

est aussi un élément du support de P. Puisque $\beta \leqslant \alpha$ par définition, on a donc $\alpha_{j+1} \leqslant \alpha_j$, d'où (1).

(2) Soient $\alpha_1, \ldots, \alpha_n$ les entiers définis récursivement dans l'énoncé, et posons alors $\alpha = (\alpha_1, \ldots, \alpha_n)$. Remarquons que, par définition même de α, on a $\alpha \in \mathrm{Supp}(P)$.

Soit $\beta = (\beta_1, \ldots, \beta_n) \in \mathrm{Supp}(P)$. Comme $a_\beta X^\beta$ est un monôme de P, on obtient $\beta_1 \leqslant \alpha_1$.

Si $\beta_1 < \alpha_1$, on a $\beta \leqslant \alpha$. Sinon, il existe $k \in [\![1, n]\!]$ tel que

$$\beta_1 = \alpha_1, \quad \ldots, \quad \beta_k = \alpha_k \quad \text{et} \quad \beta_{k+1} \neq \alpha_{k+1}.$$

Si $k = n$, $\beta = \alpha$, et si $k < n$, on a $\beta_{k+1} \leqslant \alpha_{k+1}$ par définition de α_{k+1}. Par conséquent, $\beta \leqslant \alpha$.

Ainsi, pour tout $\beta \in \mathrm{Supp}(P)$, on a $\beta \leqslant \alpha$, d'où $\alpha = D(P)$.

(3) Le point (2) montre que $\alpha_1 = \deg_{X_1}(P)$, ce qui est exactement le résultat voulu. Cela achève la démonstration. □

On continue par un lemme calculatoire.

7.18. Lemme. *Soit A un anneau commutatif et soit $n \geqslant 1$ un entier. Alors, pour tout $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ tel que $\alpha_1 \geqslant \cdots \geqslant \alpha_n$, le polynôme*

$$S = \sigma_1^{\alpha_1 - \alpha_2} \sigma_2^{\alpha_2 - \alpha_3} \cdots \sigma_{n-1}^{\alpha_{n-1} - \alpha_n} \sigma_n^{\alpha_n} \in A[X_1, \ldots, X_n]$$

est symétrique et homogène de degré $\alpha_1 + \cdots + \alpha_n$.

De plus, on a $D(S) = \alpha$, et le monôme correspondant est X^α.

Démonstration. Il est clair que S est symétrique, et homogène de degré

$$(\alpha_1 - \alpha_2) + 2(\alpha_2 - \alpha_3) + \cdots + (n-1)(\alpha_{n-1} - \alpha_n) + n\alpha_n = \alpha_1 + \cdots + \alpha_n.$$

On va maintenant calculer $D(S)$ en utilisant le lemme 7.17 (2).

On remarque que σ_j est de degré 1 en X_1, donc le degré de S en X_1 est au plus égal à

$$(\alpha_1 - \alpha_2) + (\alpha_2 - \alpha_3) + \cdots + (\alpha_{n-1} - \alpha_n) + \alpha_n = \alpha_1.$$

Or, le coefficient en X_1 de σ_j est $\sigma_{j-1}(X_2, \ldots, X_n)$ pour tout $j \in [\![2, n]\!]$. Ainsi, le coefficient de $X_1^{\alpha_1}$ de S est

$$\sigma_1(X_2, \ldots, X_n)^{\alpha_2 - \alpha_3} \cdots \sigma_{n-2}(X_2, \ldots, X_n)^{\alpha_{n-1} - \alpha_n} \sigma_{n-1}(X_2, \ldots, X_n)^{\alpha_n}.$$

Ce polynôme est non nul d'après la proposition 7.12, donc la plus grande puissance de X_1 apparaissant dans S est α_1. De plus, α_2 sera alors la plus grande puissance de X_2 intervenant dans le polynôme ci-dessus. Le même genre de raisonnement montre que le polynôme

$$\sigma_1(X_2, \ldots, X_n)^{\alpha_2 - \alpha_3} \cdots \sigma_{n-2}(X_2, \ldots, X_n)^{\alpha_{n-1} - \alpha_n} \sigma_{n-1}(X_2, \ldots, X_n)^{\alpha_n}$$

est de degré α_2 en X_2 et que le coefficient correspondant est

$$\sigma_1(X_3, \ldots, X_n)^{\alpha_3 - \alpha_4} \cdots \sigma_{n-3}(X_3, \ldots, X_n)^{\alpha_{n-1} - \alpha_n} \sigma_{n-2}(X_3, \ldots, X_n)^{\alpha_n}.$$

De proche en proche, on obtient le résultat voulu. □

On peut maintenant démontrer le théorème 7.11.

Démonstration du théorème 7.11. L'unicité d'un polynôme Q vérifiant les conditions de l'énoncé est donnée par le corollaire 7.13. Passons à l'existence de Q, que l'on va démontrer de manière algorithmique.

On commence par remarquer le fait suivant : soit $R \in A[X_1, \ldots, X_n]$ un polynôme symétrique non nul, soit $D(R) = \alpha = (\alpha_1, \ldots, \alpha_n)$ et soit $r_\alpha X^\alpha$ le monôme correspondant. Si $R - r_\alpha \sigma_1^{\alpha_1 - \alpha_2} \sigma_2^{\alpha_2 - \alpha_3} \cdots \sigma_{n-1}^{\alpha_{n-1} - \alpha_n} \sigma_n^{\alpha_n}$ est non

nul, alors

$$D(R - r_\alpha \sigma_1^{\alpha_1-\alpha_2}\sigma_2^{\alpha_2-\alpha_3}\cdots\sigma_{n-1}^{\alpha_{n-1}-\alpha_n}\sigma_n^{\alpha_n}) < D(R).$$

En effet, le lemme 7.18 montre que l'on a

$$\sigma_1^{\alpha_1-\alpha_2}\sigma_2^{\alpha_2-\alpha_3}\cdots\sigma_{n-1}^{\alpha_{n-1}-\alpha_n}\sigma_n^{\alpha_n} = X^\alpha + \sum_{\beta<\alpha} c_\beta X^\beta,$$

et ainsi

$$R - r_\alpha \sigma_1^{\alpha_1-\alpha_2}\sigma_2^{\alpha_2-\alpha_3}\cdots\sigma_{n-1}^{\alpha_{n-1}-\alpha_n}\sigma_n^{\alpha_n} = -\sum_{\beta<\alpha} r_\alpha c_\beta X^\beta,$$

d'où le fait susmentionné.

Notons également que le degré de $r_\alpha T_1^{\alpha_1-\alpha_2}T_2^{\alpha_2-\alpha_3}\cdots T_{n-1}^{\alpha_{n-1}-\alpha_n}T_n^{\alpha_n}$ est égal à α_1, qui est le degré partiel de R d'après le lemme 7.17 (3).

Soit maintenant $P \in A[X_1, \ldots, X_n]$ un polynôme symétrique, que l'on peut supposer non nul (sinon, il n'y a rien à faire).

Algorithme.–

(1) On pose $R = P$ et $Q = 0$;

(2) on calcule $D(R) = \alpha = (\alpha_1, \ldots, \alpha_n)$. On note $r_\alpha X^\alpha$ le monôme correspondant ;

(3) on effectue les opérations

$$R \longleftarrow R - r_\alpha \sigma_1^{\alpha_1-\alpha_2}\sigma_2^{\alpha_2-\alpha_3}\cdots\sigma_{n-1}^{\alpha_{n-1}-\alpha_n}\sigma_n^{\alpha_n}$$

et

$$Q \longleftarrow Q + r_\alpha T_1^{\alpha_1-\alpha_2}T_2^{\alpha_2-\alpha_3}\cdots T_{n-1}^{\alpha_{n-1}-\alpha_n}T_n^{\alpha_n}.$$

(4) si $R = 0$, on arrête. On a alors $P = Q(\sigma_1, \ldots, \sigma_n)$. Sinon, on repart à l'étape (2).

L'algorithme s'arrête nécessairement, puisque l'ensemble des éléments de $\mathbb{N}^n$ strictement inférieurs à $D(P)$ est fini et que l'algorithme nous fournit en particulier une suite strictement décroissante d'éléments de $\mathbb{N}^n$ tous inférieurs à $D(P)$. Au bout du compte, on obtient bien P comme polynôme en les polynômes symétriques élémentaires.

Supposons maintenant que l'algorithme s'arrête au bout de $k \geqslant 1$ itérations. Il existe donc des polynômes $R_1 = P, R_2, \ldots, R_k \in A[X_1, \ldots, X_n]$ symétriques et des monômes $Q_1, \ldots, Q_k \in A[Y_1, \ldots, Y_n]$ associés à des éléments de $\mathbb{N}^n$ tous distincts vérifiant

$$P = Q_1(\sigma_1, \ldots, \sigma_n) + \cdots + Q_k(\sigma_1, \ldots, \sigma_n),$$

avec $\deg(Q_i) = d_i$, où d_i est le degré partiel du polynôme symétrique R_i, pour tout $i \in [\![1, k]\!]$.

De plus, on a

$$D(R_1) > \cdots > D(R_k),$$

donc $d_1 \geqslant \cdots \geqslant d_k$, puisque le degré partiel de P_i est la première coordonnée de $D(R_i)$, par le lemme 7.17 (3). Ainsi,

$$Q = Q_1 + \cdots + Q_k$$

est une somme de k monômes tous distincts, et son degré est $\deg(Q_1) = d_1$. Autrement dit, $\deg(Q)$ est le degré partiel de P. Cela achève la démonstration. □

7.19. Remarque. Soit $P \in A[X_1, \ldots, X_n]$ un polynôme symétrique non nul.

Si P est homogène de degré d, on vérifie que les polynômes $R_1, \ldots, R_k$ obtenus sont encore homogènes de degré d.

Dans le cas général, P s'écrit comme une somme de polynômes homogènes de degrés deux à deux distincts. Comme cette décomposition est unique, et qu'un polynôme obtenu en permutant les variables dans un polynôme homogène est encore homogène de même degré, on en déduit que chacun de ces polynômes homogènes est symétrique.

Pour simplifier un peu les calculs, on a donc intérêt à appliquer l'algorithme à chacun de ces polynômes homogènes, puis à sommer les polynômes obtenus. Le résultat obtenu sera le même que si l'on avait appliqué l'algorithme directement, par unicité.

7.20. Exemples

(1) Soit $P = X_1^2X_3 + X_1^2X_2 + X_2^2X_1 + X_2^2X_3 + X_3^2X_1 + X_3^2X_2$.
Alors, P est symétrique et homogène de degré 3, et $D(P) = (2, 1, 0)$. Notons que son degré partiel est égal à 2. On pose

$$R_1 = P - \sigma_1^{2-1}\sigma_2^{1-0}\sigma_3^0 = P - \sigma_1\sigma_2 \quad \text{et} \quad Q_1 = T_1T_2.$$

On vérifie que l'on a $R_1 = -3X_1X_2X_3$. Évidemment, on reconnaît immédiatement le polynôme $-3\sigma_3$, mais on va quand même appliquer l'algorithme pour voir que cela fonctionne bien.
On a $D(R_1) = (1, 1, 1)$. On pose donc

$$R_2 = R_1 - (-3\sigma_1^{1-1}\sigma_2^{1-1}\sigma_3^1) = R_1 + 3\sigma_3 \quad \text{et} \quad Q_2 = -3T_3.$$

On a alors $R_3 = 0$.
Ainsi, $P = \sigma_1\sigma_2 - 3\sigma_3$ et $Q = T_1T_2 - 3T_3$, qui est bien de degré 2.

(2) Soit $P = X_1^2 + X_2^2 + X_3^2$.
Alors, P est symétrique et homogène de degré 2, et $D(P) = (2, 0, 0)$.

On pose

$$R_1 = P - \sigma_1^{2-0}\sigma_2^{0-0}\sigma_3^0 = P - \sigma_1^2 \quad \text{et} \quad Q_1 = T_1^2.$$

On vérifie que l'on a

$$R_1 = -2(X_1X_2 + X_1X_3 + X_2X_3).$$

On a donc $R_2 = R_1 + 2\sigma_2 = 0$ et $Q_2 = -2T_2$.

Finalement, $P = \sigma_1^2 - 2\sigma_2$ et $Q = T_1^2 - 2T_2$.

(3) Soit $P = X_1^2+X_2^2+X_3^2+X_1^2X_3+X_1^2X_2+X_2^2X_1+X_2^2X_3+X_3^2X_1+X_3^2X_2$.

Le polynôme P est la somme d'un polynôme homogène de degré 2, et d'un polynôme homogène de degré 3, qui ne sont rien d'autre que les polynômes des exemples précédents. On a donc immédiatement

$$P = Q(\sigma_1, \sigma_2, \sigma_3), \quad \text{où } Q = (T_1^2 - 2T_2) + (T_1T_2 - 3T_3).$$

Nous allons maintenant appliquer l'algorithme directement au polynôme P, et retrouver le résultat précédent.

On a $D(P) = (2, 1, 0)$. On pose donc

$$R_1 = P - \sigma_1\sigma_2 \quad \text{et} \quad Q_1 = T_1T_2.$$

On a alors

$$R_1 = X_1^2 + X_2^2 + X_3^2 - 3X_1X_2X_3.$$

Comme $D(R_1) = (2, 0, 0)$, on pose cette fois

$$R_2 = R_1 - \sigma_1^2 \quad \text{et} \quad Q_2 = T_1^2.$$

On a donc

$$R_2 = -2(X_1X_2 + X_1X_3 + X_2X_3) - 3X_1X_2X_3.$$

On a ensuite $D(R_2) = (1, 1, 1)$, puis

$$R_3 = R_2 + 3\sigma_3 \quad \text{et} \quad Q_3 = -3T_3.$$

Comme $R_3 = -2(X_1X_2 + X_1X_3 + X_2X_3)$, on a $Q_4 = -2T_2$ et ainsi

$$Q = T_1T_2 + T_1^2 - 3T_3 - 2T_2.$$

(4) Exprimons le polynôme symétrique P

$$X_1^2(X_2X_3 + X_2X_4 + X_3X_4) + X_2^2(X_1X_3 + X_1X_4 + X_3X_4)$$
$$+ X_3^2(X_1X_2 + X_1X_4 + X_2X_4) + X_4^2(X_1X_2 + X_1X_3 + X_2X_3)$$

en termes de $\sigma_1, \ldots, \sigma_4$.

C'est un polynôme homogène de degré 4, et l'on a $D(P) = (2, 1, 1, 0)$.

On pose alors $R_1 = P - \sigma_1^{2-1}\sigma_2^{1-1}\sigma_3^1\sigma_4^0 = P - \sigma_1\sigma_3$ et $Q_1 = T_1T_3$. Un simple calcul montre que l'on a

$$R_1 = -4X_1X_2X_3X_4 = -4\sigma_4.$$

Par conséquent, $P = \sigma_1\sigma_3 - 4\sigma_4$, et $Q = T_1T_3 - 4T_4$.

Nous donnerons plusieurs applications de ce théorème de structure lorsque nous aborderons la théorie des corps.

8. Exercices

E1. Démontrer la proposition 1.4.

E2. Détailler les exemples 1.3 (1) – (5).

E3. Démontrer qu'un sous-anneau d'un anneau intègre est intègre.

E4. Les ensembles suivants, pour les lois considérées, sont-ils des anneaux ?

$(\mathbb{Q},+,\cdot)$; $(\mathbb{Q},-,\cdot)$; $(\mathbb{Q},+,:)$; $(\mathbb{R}\setminus\mathbb{Q},+,\cdot)$; $(\mathcal{F}(\mathbb{R},\mathbb{R}),+,\cdot)$; $(\mathcal{F}(\mathbb{R},\mathbb{R}),+,\circ)$;

$(C^0(I,\mathbb{R}),+,\cdot)$, où I est un intervalle de $\mathbb{R}$; $(\mathbb{Z}[\sqrt{2}],+,\cdot)$.

E5. Montrer que $3\mathbb{Z}/6\mathbb{Z}$ est un anneau pour les lois de $\mathbb{Z}/6\mathbb{Z}$. Est-ce un sous-anneau de $\mathbb{Z}/6\mathbb{Z}$? Généraliser.

E6. Les sous-ensembles suivants sont-ils des sous-anneaux de $\mathbb{Z}[X]$?

(1) $\{P \in \mathbb{Z}[X] \mid \deg(P) \leqslant n\}$ $(n \geqslant 0)$

(2) $\{P \in \mathbb{Z}[X] \mid \deg(P) \equiv 0 \mod 2\} \cup \{0\}$

(3) $\{XQ \mid Q \in \mathbb{Z}[X]\}$

(4) $\{Q(X^2) \mid Q \in \mathbb{Z}[X]\}$

E7. Détailler les exemples 1.7.

E8. Parmi les anneaux suivants, lesquels sont intègres ?

$$\mathbb{Q}, \quad \mathbb{Z}[\sqrt{2}], \quad \mathbb{Z}/6\mathbb{Z}, \quad \mathbb{Z}/7\mathbb{Z}, \quad C^0(\mathbb{R},\mathbb{R}).$$

E9. Soit A un anneau. Quel est le plus petit sous-anneau de A ?

E10. Donner toutes les tables d'addition et de multiplication possibles des anneaux à 2, 3 ou 4 éléments.

Indication. Remarquer qu'un anneau A contient toujours au moins 0_A et 1_A. Pour le cas d'un anneau A à 4 éléments, remarquer que $(A,+)$ est soit cyclique, soit isomorphe à $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}$.

E11. Soit $d \in \mathbb{Z}$ sans facteurs carrés (en particulier, $d \neq 1$).

a. Montrer que $\mathbb{Z}[\sqrt{d}] = \{a+b\sqrt{d} \mid a, b \in \mathbb{Z}\}$ est un anneau. Montrer que l'écriture d'un élément de $\mathbb{Z}[\sqrt{d}]$ sous la forme $a + b\sqrt{d}$ est unique.

b. On suppose que $d \equiv 1 \mod 4$. Montrer que

$$\mathbb{Z}\Big[\frac{-1+\sqrt{d}}{2}\Big] = \Big\{a + b\frac{-1+\sqrt{d}}{2} \;\Big|\; a, b \in \mathbb{Z}\Big\}$$

est un anneau. Est-ce encore vrai si $d \not\equiv 1 \mod 4$? Montrer que l'écriture d'un élément de $\mathbb{Z}\Big[\frac{-1+\sqrt{d}}{2}\Big]$ sous la forme $a + b\frac{-1+\sqrt{d}}{2}$ est unique.

c. Les anneaux $\mathbb{Z}[\sqrt{d}]$ et $\mathbb{Z}\Big[\frac{-1+\sqrt{d}}{2}\Big]$ sont-ils intègres ? des corps ?

d. Montrer que $\mathbb{Q}[\sqrt{d}] = \{a + b\sqrt{d} \mid a, b \in \mathbb{Q}\}$ est un corps.

e. Montrer que $\mathbb{Z}[\sqrt{d}]$ est le plus petit sous-anneau de $\mathbb{C}$ contenant $\sqrt{d}$.

f. Montrer que $\mathbb{Q}[\sqrt{d}]$ est le plus petit sous-corps de $\mathbb{C}$ contenant $\sqrt{d}$.

E12. Soit B un anneau commutatif, soit $\omega \in B$ et soit A un sous-anneau de B. On note $A[\omega] = \{P(\omega) \mid P \in A[X]\}$.

a. Montrer que $A[\omega]$ est le plus petit sous-anneau de B contenant à la fois A et ω.

b. On suppose que $A = \mathbb{Z}$ et $B = \mathbb{C}$. Montrer que $\mathbb{Z}[\omega]$ est le plus petit sous-anneau de $\mathbb{C}$ contenant ω.

c. On suppose que $A = \mathbb{Z}$, et qu'il existe un polynôme $P_0 \in \mathbb{Z}[X]$ **unitaire** de degré $d \geqslant 1$ tel que $P_0(\omega) = 0$. Montrer que l'on a

$$\mathbb{Z}[\omega] = \{a_{d-1}\omega^{d-1} + \cdots + a_1\omega + a_0 \mid a_i \in \mathbb{Z}\}.$$

Est-ce vrai si f n'est pas unitaire ?

E13. On dit qu'un anneau A est *un anneau de Boole* si pour tout $a \in A$, on a $a^2 = a$.

a. Soit E un ensemble. On note $\mathscr{P}(E)$ l'ensemble des parties de E. Pour tous A, $B \in \mathscr{P}(E)$, on pose

$$A\Delta B = (A \cup B) \setminus (A \cap B).$$

(i) Montrer que l'ensemble $\mathscr{F}(E, \mathbb{Z}/2\mathbb{Z})$ des fonctions de E dans $\mathbb{Z}/2\mathbb{Z}$ est un anneau de Boole.

(ii) Montrer que l'application

$$\begin{array}{rcl} \varphi : \mathscr{F}(E, \mathbb{Z}/2\mathbb{Z}) & \longrightarrow & \mathscr{P}(E) \\ f & \longmapsto & f^{-1}(\{\overline{1}\}) \end{array}$$

est bijective, et expliciter son inverse.

(iii) Pour tous f, $g \in \mathscr{F}(E, \mathbb{Z}/2\mathbb{Z})$, calculer $\varphi(f+g)$ et $\varphi(fg)$ en fonction de $\varphi(f)$ et $\varphi(g)$.

(iv) En déduire que $(\mathscr{P}(E), \Delta, \cap)$ est un anneau de Boole. On précisera quelle est la loi d'addition, et quelle est la loi de multiplication, ainsi que les éléments neutres.

(v) Montrer que l'on a un isomorphisme d'anneaux

$$\mathscr{F}(E, \mathbb{Z}/2\mathbb{Z}) \simeq \mathscr{P}(E).$$

b. Montrer qu'un sous-anneau d'un anneau de Boole est un anneau de Boole.

c. Montrer que l'ensemble A des suites à valeurs dans $\mathbb{Z}/2\mathbb{Z}$ qui sont constantes à partir d'un certain rang est un anneau de Boole infini dénombrable. Est-il isomorphe à $\mathscr{P}(E)$ pour un certain ensemble E ?

d. Soit A un anneau de Boole. Montrer que $2a = 0$ pour tout $a \in A$, et que A est commutatif.

e. Quels sont les anneaux de Boole intègres ?

f. Montrer qu'un anneau de Boole n'a pas d'éléments nilpotents non nuls.

g. Montrer que, dans un anneau de Boole, 1 est le seul élément inversible.

E14. Soit A un anneau fini.

a. Montrer qu'un élément $a \in A$ est soit nul, soit inversible, soit diviseur de zéro. En particulier, un anneau intègre fini est un corps [7].

Indication. Considérer les applications de multiplication à gauche et à droite par a.

b. Est-ce vrai si A est infini ?

7. Ce résultat explique ici les résultats de la proposition 3.9.

E15. Soit $A = \mathbb{Z}[\sqrt{d}]$. Pour tout $z = a + b\sqrt{d} \in A$, on pose $\overline{z} = a - b\sqrt{d}$, ainsi que $N(z) = z\overline{z} = a^2 - db^2$.

a. Vérifier que, pour tout $z \in A$, on a $\overline{z} \in A$ et $N(z) \in \mathbb{Z}$, et que, pour tous z, $z' \in A$, on a

$$N(zz') = N(z) \times N(z').$$

b. En déduire que $z \in A^\times \iff N(z) = \pm 1$.

c. On suppose que $d < 0$.

(i) Montrer que $z \in A^\times \iff N(z) = 1$.

(ii) On suppose que $d \neq -1$. Montrer que $A^\times = \{\pm 1\}$.

(iii) On suppose que $d = -1$. Montrer que $A^\times = \{\pm 1, \pm i\}$.

d. On suppose que $d > 0$. Soit $W = \{z \in A^\times \mid z > 1\}$. On admet que l'ensemble W est non vide. On veut montrer dans cette question que W admet un plus petit élément.

(i) Soit $z = a + b\sqrt{d} \in W$. Montrer que $-1 < a - b\sqrt{d} < 1$. En déduire que a, $b \geqslant 1$.

(ii) Soit $\zeta \in W$. On suppose que $z = a + b\sqrt{d} \in W$ vérifie $z \leqslant \zeta$. Montrer que

$$1 \leqslant a, b \leqslant \zeta.$$

En déduire que l'ensemble $\{z \in W \mid z \leqslant \zeta\}$ est fini non vide, et admet un plus petit élément ω.

(iii) Montrer que ω est le plus petit élément de W.

On veut montrer que $A^\times = \{\pm\omega^n \mid n \in \mathbb{Z}\}$. Soit $z \in A^\times$.

(i) On suppose ici que $z > 0$. Montrer qu'il existe $n \in \mathbb{Z}$ tel que

$$\omega^{n-1} < z \leqslant \omega^n.$$

Montrer que $\dfrac{z}{\omega^{n-1}} \in W$. En déduire que $z = \omega^n$.

(ii) On suppose ici que $z < 0$. Montrer que $-z = \omega^n$ pour un certain entier $n \in \mathbb{Z}$.

(iii) Montrer que l'application

$$\begin{aligned} \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z} &\longrightarrow A^\times \\ (\overline{m}, n) &\longmapsto (-1)^m \omega^n \end{aligned}$$

est un isomorphisme de groupes.

e. Déterminer les éléments inversibles de $\mathbb{Z}[\sqrt{2}]$ et de $\mathbb{Z}[\sqrt{7}]$.

E16. On se propose d'étudier dans cet exercice la structure de $(\mathbb{Z}/p^r\mathbb{Z})^\times$, où p est un nombre premier et $r \geqslant 1$.

a. Montrer que $|(\mathbb{Z}/p^r\mathbb{Z})^\times| = p^{r-1}(p-1)$.

b. On suppose que p est impair. Soit $r \geqslant 1$. On rappelle que $(\mathbb{Z}/p\mathbb{Z})^\times$ est cyclique.

(i) Soit $m \in \mathbb{Z}$ tel que la classe de m modulo p engendre $(\mathbb{Z}/p\mathbb{Z})^\times$. Montrer que $\overline{m}^{p^{r-1}} \in (\mathbb{Z}/p^r\mathbb{Z})^\times$ est d'ordre $p-1$.

(ii) Montrer par récurrence que, pour tout $k \geqslant 0$, il existe $u_k \in \mathbb{Z}, p \nmid u_k$ tel que

$$(1+p)^{p^k} = 1 + u_k p^{k+1}.$$

En déduire que $\overline{1+p} \in (\mathbb{Z}/p^r\mathbb{Z})^\times$ est d'ordre p^{r-1}.

(iii) Déduire des questions précédentes que le groupe $(\mathbb{Z}/p^r\mathbb{Z})^\times$ est cyclique d'ordre $p^{r-1}(p-1)$.

c. On suppose que $p = 2$.

(i) Montrer que $(\mathbb{Z}/2^r\mathbb{Z})^\times$ est cyclique pour $r = 1, 2$. Est-ce encore vrai pour $r = 3$?
On suppose dans la suite que $r \geqslant 3$.

(ii) Montrer par récurrence que, pour tout $k \geqslant 0$, il existe $u_k \in \mathbb{Z}$ impair tel que

$$5^{2^k} = 1 + u_k 2^{k+2}.$$

En déduire que $\overline{5}$ est d'ordre 2^{r-2}.

(iii) Montrer que $(\mathbb{Z}/2^r\mathbb{Z})^\times = \langle\overline{-1}\rangle \odot \langle\overline{5}\rangle$.

(iv) En déduire que $(\mathbb{Z}/2^r\mathbb{Z})^\times \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{r-2}\mathbb{Z}$, pour $r \geqslant 3$.

d. Résumer les résultats obtenus.

E17. Si $P \in \mathbb{C}[X]$ est non nul, on note $\nu(P)$ le nombre de racines distinctes de P et $D(P) = \operatorname{pgcd}(P, P')$.

a. Montrer que $\deg\big(D(P)\big) = \deg(P) - \nu(P)$.

b. On suppose que A, B, $C \in \mathbb{C}[X]$ sont des polynômes premiers entre eux deux à deux non constants, tels que $A + B + C = 0$.

(1) Vérifier que $A'B - AB' = AC' - A'C$.

(2) Montrer que $D(A)D(B)D(C)$ divise $A'B - AB'$.

(3) En déduire que

$$\max\big(\deg(A), \deg(B), \deg(C)\big) \leqslant \nu(A) + \nu(B) + \nu(C) - 1.$$

c. Montrer que pour tout $n \geqslant 3$, il n'existe pas de polynômes P, Q, R de $\mathbb{C}[X]$ premiers entre eux deux à deux non constants tels que l'on a $P^n + Q^n = R^n$.

E18. Soit A un anneau commutatif et soit G un groupe quelconque. On note $A[G]$ l'ensemble des familles $(a_{g'})_{g' \in G}$, où les $a_g \in A$ sont presque tous nuls.

Pour tout $g \in G$, on note $e_g \in A[G]$ l'élément défini par

$$e_g = (\delta_{g,g'})_{g' \in G}.$$

Pour tout $a \in A$, et tout $x = (a_{g'})_{g' \in G} \in A[G]$, on note

$$a \cdot x = (a a_{g'})_{g' \in G}.$$

a. Montrer que $(A[G], +)$ est un groupe abélien.

b. Montrer que tout élément $x \in A[G]$ s'écrit de manière unique sous la forme

$$x = \sum_{g \in G} a_g \cdot e_g,$$

où les a_g sont presque tous nuls.

On définit une multiplication sur $A[G]$ comme suit :

$$\Big(\sum_{g \in G} a_g \cdot e_g\Big)\Big(\sum_{g \in G} b_g \cdot e_g\Big) = \sum_{g \in G}\Big(\sum_{h \in G} a_h b_{h^{-1}g}\Big) \cdot e_g.$$

c. Montrer que $A[G]$, muni de l'addition, et de la multiplication précédente, est un anneau commutatif.

d. Montrer que pour tout morphisme de groupes

$$f : G \longrightarrow B^\times,$$

où B est un anneau commutatif, il existe un unique morphisme d'anneaux $\varphi : \mathbb{Z}[G] \longrightarrow B$ tel que

$$\varphi(e_g) = f(g), \quad \text{pour tout } g \in G.$$

E19. Soit A un anneau commutatif. Une *fonction arithmétique à valeurs dans A* est une fonction $f : \mathbb{N}^* \longrightarrow A$. On munit l'ensemble $\mathscr{A}$ des fonctions arithmétiques à valeurs dans A de l'addition des fonctions. Pour tous f, $g \in \mathscr{A}$, on note $f * g$ la fonction arithmétique définie par

$$(f * g)(n) = \sum_{d \mid n} f(d)\, g\Big(\frac{n}{d}\Big), \quad \text{pour tout } n \geqslant 1,$$

où la somme est prise sur tous les diviseurs d positifs de n.

a. Montrer que $(\mathscr{A}, +, *)$ est un anneau, dont le neutre pour la loi $*$ est la fonction e définie par

$$e(n) = \delta_{n,1} 1_A, \quad \text{pour tout } n \geqslant 1.$$

b. On considère la fonction $\mu \in \mathscr{A}$ définie par

$$\mu(n) = \begin{cases} (-1)^r 1_A, & \text{si } n \text{ est produit de } r \text{ nombres premiers distincts} \\ 0_A, & \text{s'il existe un nombre premier } p \text{ tel que } p^2 \mid n. \end{cases}$$

On note $\mathbf{1}$ la fonction constante égale à 1. Calculer $\mu * \mathbf{1}$.

c. En déduire que si $f, g \in \mathscr{A}$ vérifie $g(n) = \sum_{d|n} f(d)$ pour tout $n \geqslant 1$, alors on a

$$f(n) = \sum_{d|n} \mu(d) g\Big(\frac{n}{d}\Big).$$

d. Montrer que $f \in \mathscr{A}^\times$ si, et seulement si, $f(1) \in A^\times$.

On dit que $f \in \mathscr{A}$ est *multiplicative* si $f(1) = 1_A$, et si pour tous entiers $m, n \geqslant 1$ premiers entre eux, on a $f(mn) = f(m)f(n)$.

e. Montrer que si f et g sont multiplicatives, $f * g$ est multiplicative. On pourra d'abord démontrer que si m et n sont premiers entre eux, alors tout diviseur d de mn s'écrit de manière unique sous la forme $d = d_1 d_2$, où $d_1 \mid m$ et $d_2 \mid n$.

f. Montrer qu'une fonction multiplicative est inversible, et que son inverse est une fonction multiplicative.

E20. Soit $\mathbb{H} \subset \mathrm{M}_2(\mathbb{C})$ l'ensemble défini par

$$\mathbb{H} = \left\{ \begin{pmatrix} z_1 & -z_2 \\ \overline{z}_2 & \overline{z}_1 \end{pmatrix} \;\middle|\; z_1, z_2 \in \mathbb{C} \right\}.$$

a. Montrer que $\mathbb{H}$ est un sous-anneau non trivial de $\mathrm{M}_2(\mathbb{C})$.

b. Montrer que $\mathbb{H}$ n'est pas commutatif.

c. Montrer que $\mathbb{H}$ est un anneau à division.

Indication. Si $M \in \mathbb{H}$ a un inverse dans $\mathbb{H}$, il a un inverse dans $\mathrm{M}_2(\mathbb{C})$. Quel est-il ? Vérifier alors que cet inverse dans $\mathrm{M}_2(\mathbb{C})$ appartient à $\mathbb{H}$.

d. Montrer que le polynôme $X^2 + 1_{\mathbb{H}} \in \mathbb{H}[X]$ a une infinité de racines.

E21. Soit A un anneau. Montrer que l'application

$$\begin{aligned} \Theta_A \colon \mathbb{Z} &\longrightarrow A \\ m &\longmapsto m \cdot 1_A \end{aligned}$$

est un morphisme d'anneaux.

E22.

a. Montrer qu'un anneau à n éléments ($n \geqslant 1$) et de caractéristique n est isomorphe à $\mathbb{Z}/n\mathbb{Z}$.

b. Soit p un nombre premier. Montrer qu'un anneau à p éléments est isomorphe à $\mathbb{Z}/p\mathbb{Z}$.

c. Calculer la caractéristique des anneaux suivants :

$$\mathbb{Z}, \quad \mathbb{Z}/n\mathbb{Z}, \quad \mathbb{Q} \times \mathbb{Z}/3\mathbb{Z}, \quad \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/10\mathbb{Z},$$
$$A_1 \times \cdots \times A_n, \quad A_i \text{ anneau.}$$

E23. Soit A un anneau intègre. Montrer que l'application

$$\begin{aligned} A[X] &\longrightarrow \mathscr{F}(A, A) \\ P &\longmapsto \tilde{P} \end{aligned}$$

est injective si, et seulement si, A est infini.

E24. Les sous-ensembles suivants de $\mathbb{Z}[X]$ sont-ils des sous-anneaux ? Des idéaux ?

(1) $\{P \in \mathbb{Z}[X] \mid P(1) = 1\}$

(2) $\{P \in \mathbb{Z}[X] \mid P(1) = 0\}$

(3) $\{P \in \mathbb{Z}[X] \mid P(1) \in 3\mathbb{Z}\}$

(4) $\{P \in \mathbb{Z}[X] \mid P'(1) = 1\}$

E25. Les applications

$$\begin{aligned} \mathbb{Z}[X] &\longrightarrow \mathbb{Z} \\ P &\longmapsto P(1) \end{aligned} \qquad \text{et} \qquad \begin{aligned} \mathbb{Z}[X] &\longrightarrow \mathbb{Z} \\ P &\longmapsto P'(0) \end{aligned}$$

sont-elles des morphismes d'anneaux ?

E26. Trouver une condition nécessaire et suffisante pour qu'il existe un morphisme d'anneaux $\mathbb{Z}/m\mathbb{Z} \longrightarrow \mathbb{Z}/n\mathbb{Z}$.

E27. Soit K un corps. On veut montrer que les groupes $(K, +)$ et $(K^\times, \cdot)$ ne sont jamais isomorphes. On procède par l'absurde, et l'on suppose qu'il existe un isomorphisme $\varphi : K^\times \xrightarrow{\sim} K$.

a. Montrer que $2 \cdot 1_K = 0_K$.

Indication. Si $1_K \neq -1_K$, trouver une contradiction en remarquant que $(-1_K)^2 = 1_K$.

b. En déduire que pour tout $y \in K$, on a $2 \cdot y = 0_K$, puis que $x^2 = 1_K$ pour tout $x \in K^\times$.

c. Montrer alors que K a au plus deux éléments, et conclure.

E28. Démontrer le lemme 3.14 et la proposition 3.15.

E29. Soit A un anneau intègre. Montrer que si A est un corps, l'application

$$\begin{aligned} \iota_A \colon A &\longrightarrow K_A \\ a &\longmapsto \frac{a}{1} \end{aligned}$$

est un isomorphisme d'anneaux.

E30. Soit K un corps.

a. Montrer que les seuls idéaux bilatères de $\mathrm{M}_2(K)$ sont (0) et $\mathrm{M}_2(K)$.

Indication. Montrer que si $\mathfrak{a}$ est un idéal bilatère non nul de $\mathrm{M}_2(K)$, alors il contient les matrices $\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ et $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$. Pour cela, on multipliera un élément non nul de $\mathfrak{a}$ par des matrices bien choisies.

b. Soit A un anneau commutatif. Montrer que les idéaux bilatères de l'anneau $\mathrm{M}_2(A)$ sont de la forme $\mathrm{M}_2(\mathfrak{a})$, où $\mathfrak{a}$ est un idéal bilatère de A.

E31. Exhiber un idempotent non trivial de $\mathrm{M}_2(K)$. Montrer que néanmoins, il n'existe pas d'anneaux A_1, A_2 tous deux non triviaux tels que $\mathrm{M}_2(K) \simeq A_1 \times A_2$.

E32. On appelle *pseudo-anneau* un ensemble A muni de deux lois « $+$ » et « $\cdot$ », et vérifiant tous les axiomes définissant un anneau, sauf l'existence d'un élément neutre pour la loi produit. On définit un morphisme de pseudo-anneau de manière évidente.

a. Soit A un anneau. Montrer que tout idéal bilatère de A est un pseudo-anneau.

b. Soit A un pseudo-anneau. Pour tous n, $n' \in \mathbb{Z}$ et tous a, $a' \in A$, on pose

$$(n, a) \cdot (n', a') = (nn', na' + n'a + aa').$$

(i) Montrer que le groupe abélien $(A', +)$, muni de la loi produit précédente, est un anneau, dont le neutre pour la loi produit est $(1, 0_A)$.

(ii) Montrer que l'application

$$\begin{aligned} A &\longrightarrow A' \\ a &\longmapsto (0, a) \end{aligned}$$

est un morphisme de pseudo-anneaux injectif, et montrer que son image est un idéal bilatère de A'.

c. On suppose que A est un anneau. Montrer que l'application

$$\begin{aligned} \varphi : \mathbb{Z} \times A &\longrightarrow A' \\ (n, a) &\longmapsto (n, a - n \cdot 1_A) \end{aligned}$$

est un isomorphisme d'anneaux.

d. Soit $(G, +)$ un groupe abélien. Montrer que l'on munit G d'une structure de pseudo-anneau en posant

$$gg' = 0_G, \quad \text{pour tous } g, g' \in G.$$

e. En déduire que, pour tout groupe abélien $(G, +)$, il existe un anneau commutatif A' et un morphisme injectif $G \longrightarrow A'$ de groupes abéliens.

E33. Démontrer le lemme 6.24.

E34. Soit $f : A \longrightarrow B$ un morphisme d'anneaux.

a. On suppose que f est surjectif. Montrer que l'image d'un idéal à gauche (resp. à droite, resp. bilatère) de A par f est un idéal à gauche (resp. à droite, resp. bilatère) de B.

b. Donner un contre-exemple lorsque f n'est pas surjectif.

E35. Soit $f : A \longrightarrow B$ un morphisme d'anneaux. On suppose que A est un anneau à division et que B est non trivial. Montrer que f est injectif. Que se passe-t-il si B est trivial ?

E36. Soit $\mathfrak{a} = (2, X)$ l'idéal de $\mathbb{Z}[X]$ engendré par 2 et X.

a. Montrer que $\mathfrak{a} = \{P \in \mathbb{Z}[X] \mid P(0) \text{ est pair }\}$.

b. Montrer que $\mathfrak{a}$ n'est pas un idéal principal.

Indication. Supposer que $\mathfrak{a} = (P)$. En considérant le fait que $2 \in \mathfrak{a}$, montrer que $P = a_0$, $a_0 \in \mathbb{Z}$. Montrer que $a_0 \mid 2$, puis conclure en remarquant que $X \in \mathfrak{a}$.

E37. Soient A et B deux anneaux.

a. On suppose que A et B sont commutatifs.

Montrer que pour tout morphisme d'anneaux $f : A \longrightarrow B$, tout $n \geqslant 1$ et tout $\mathbf{b} = (b_1, \ldots, b_n) \in B^n$, il existe un unique morphisme d'anneaux

$$\varphi_{f,\mathbf{b}} : A[X_1, \ldots, X_n] \longrightarrow B$$

tel que

$$\varphi_{f,\mathbf{b}|_A} = f \quad \text{et} \quad \varphi_{f,\mathbf{b}}(X_i) = b_i, \quad \text{pour tout } i \in [\![1, n]\!].$$

Montrer que tout morphisme d'anneaux $A[X_1, \ldots, X_n] \longrightarrow B$ est de cette forme.

b. Soit K un corps, et soit B une K-algèbre non nécessairement commutative. Soient $b_1, \ldots, b_n \in B$ des éléments commutant deux à deux. Montrer qu'il existe un unique morphisme de K-algèbres

$$ev_{b_1,\ldots,b_n} : K[X_1, \ldots, X_n] \longrightarrow B$$

qui envoie X_i sur b_i. Expliciter l'image de $P \in K[X_1, \ldots, X_n]$ par ce morphisme en fonction de $b_1, \ldots, b_n$ et des coefficients de P.

Par analogie avec le cas des polynômes en une indéterminée, l'image de P par le morphisme précédent est notée $P(b_1, \ldots, b_n)$.

Ce morphisme sera utilisé dans l'exercice suivant.

E38. Soit K un corps. Soient n, $m \geqslant 1$ et soit $M = (M_{ij})_{i,j} \in \mathrm{M}_{nm}(K)$ une matrice par blocs, où les matrices $M_{ij} \in \mathrm{M}_m(K)$ commutent deux à deux. On se propose de démontrer l'égalité

$$\det(M) = \det\Big(\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) M_{\sigma(1)1} \cdots M_{\sigma(n)n}\Big).$$

La matrice $X = (X_{ij})_{i,j}$ à coefficients dans $K(X_{11}, \ldots, X_{nn})$ a un déterminant appartenant à $K[X_{11}, \ldots, X_{nn}]$.

On pose donc

$$\Delta = \det\big((X_{ij})_{i,j}\big) \in K[X_{11}, \ldots, X_{nn}].$$

On note aussi $\Delta_{11}, \ldots, \Delta_{nn}$ les mineurs de la matrice X.

La formule étant trivialement vraie pour $n = 1$, on va procéder par récurrence sur n, et l'on suppose la formule vraie pour tout corps K et toute famille de $(n-1)^2$ matrices de $\mathrm{M}_m(K)$ commutant deux à deux, pour un certain entier $n \geqslant 2$.

Soient K et un corps et $M_{11}, \ldots, M_{nn} \in \mathrm{M}_m(K)$ n^2 matrices commutant deux à deux à deux. Soit T une nouvelle indéterminée sur K. Pour tous i, $j \in n\,[\![1, n]\!]$, on pose

$$N_{ij} = M_{ij} + Z\delta_{ij}\mathrm{I}_m.$$

Les matrices $N_{11}, \ldots, N_{nn}$ commutant deux à deux, d'après l'exercice

précédent, on peut alors définir

$$N'_{ij} = (-1)^{i+j}\Delta_{ij}(N_{11},\dots,N_{nn}) \in \mathrm{M}_n\big(K(T)\big).$$

a. Vérifier que $\sum_{i=1}^{n}(-1)^{k+i}X_{kji}\Delta_{ki} = \delta_{jk}\Delta$, pour tout $j \in [\![1,n]\!]$.

b. Soit $N = (N_{ij})_{i,j}$, et soit

$$U = \begin{pmatrix} N'_{11} & & \\ N'_{12} & \mathrm{I}_m & \\ \vdots & \vdots & \\ N'_{1n} & 0 & \mathrm{I}_m \end{pmatrix}.$$

Calculer le produit NU, et en déduire l'égalité

$$\det(N)\det(U) = \det\big(\Delta(N_{11},\dots,N_{nn})\big)\det(Q),$$

où $Q = (N_{ij})_{i,j\in[\![2,n]\!]}$.

c. En utilisant l'hypothèse de récurrence, montrer les égalités

$$\det(Q) = \det(N'_{11}) = \det(U).$$

d. Montrer que $\det(Q)$ est un polynôme unitaire de $K[T]$.

e. En déduire que $\det(N) = \det\big(\Delta(N_{11},\dots,N_{nn})\big) \in K[T]$. Comparer les termes constants, et conclure[8].

E39. Cet exercice nécessite de connaître le théorème de structure des polynômes symétriques et l'algorithme de calcul qui va avec.

Écrire les polynômes $X_1^3 + X_2^3 + X_3^3$ et $X_1^4 + \cdots + X_4^4$ comme polynômes en les polynômes symétriques élémentaires de $\mathbb{Z}[X_1, X_2, X_3]$ et de $\mathbb{Z}[X_1, X_2, X_3, X_4]$ respectivement.

8. Le résultat de cet exercice est tiré de la démonstration du lemme 1 de [15, Chapitre III, §9].

Chapitre XVI

Anneaux quotients, théorème chinois

1. Anneaux quotients, théorème de factorisation

Rappelons que l'on a muni le groupe additif $\mathbb{Z}/n\mathbb{Z}$ d'une structure d'anneau en posant

$$\overline{a}\ \overline{b} = \overline{ab}.$$

Il est facile de voir alors que l'application

$$\begin{aligned} \pi : \mathbb{Z} &\longrightarrow \mathbb{Z}/n\mathbb{Z} \\ a &\longmapsto \overline{a} \end{aligned}$$

est un morphisme d'anneaux.

On se propose maintenant de généraliser cette construction à un anneau arbitraire. Soit A un anneau, et soit $\mathfrak{a}$ un sous-groupe additif de A. On veut maintenant munir $A/\mathfrak{a}$ d'une structure d'anneau telle que la projection canonique $\pi : A \longrightarrow A/\mathfrak{a}$ soit un morphisme d'anneaux.

Si une telle structure d'anneau existe, alors pour tous $\overline{a}, \overline{b} \in A/\mathfrak{a}$, on doit avoir

$$\overline{a} + \overline{b} = \pi(a) + \pi(b) = \pi(a+b) = \overline{a+b},$$

ainsi que

$$\overline{a}\ \overline{b} = \pi(a)\pi(b) = \pi(ab) = \overline{ab}.$$

Puisque $\mathfrak{a}$ est un sous-groupe additif du groupe abélien $(A, +)$, $\mathfrak{a}$ est distingué dans A. Ainsi, la loi d'addition précédente est bien définie, et $(A/\mathfrak{a}, +)$ est un groupe abélien.

Examinons le cas de la loi produit. Dans le cas où elle est bien définie, pour tout $x \in \mathfrak{a}$ et tout $a \in A$, on doit avoir

$$\overline{ax} = \overline{a}\ \overline{x} = \overline{a}\ \overline{0} = \overline{a0} = \overline{0}.$$

Autrement dit, on a $ax \in \mathfrak{a}$. On obtient de même que $xa \in \mathfrak{a}$. Ainsi, si la loi produit est bien définie, $\mathfrak{a}$ est nécessairement un idéal bilatère de A.

La proposition suivante montre que réciproquement, si $\mathfrak{a}$ est un idéal bilatère de A, ces deux lois sont bien définies, et répondent à notre problème.

1.1. Proposition. *Soit A un anneau, et soit $\mathfrak{a}$ un idéal bilatère de A. Les applications*

$$\begin{aligned} A/\mathfrak{a} \times A/\mathfrak{a} &\longrightarrow A/\mathfrak{a} \\ (\overline{a}, \overline{b}) &\longmapsto \overline{a+b} \end{aligned}$$

et

$$\begin{aligned} A/\mathfrak{a} \times A/\mathfrak{a} &\longrightarrow A/\mathfrak{a} \\ (\overline{a}, \overline{b}) &\longmapsto \overline{ab} \end{aligned}$$

sont bien définies, et confèrent à $A/\mathfrak{a}$ une structure d'anneau, de neutre additif $\overline{0}$, et de neutre multiplicatif $\overline{1}$.

De plus, l'application

$$\pi : A \longrightarrow A/\mathfrak{a}$$

est un morphisme d'anneaux.

Démonstration. Comme $\mathfrak{a}$ est un sous-groupe de $(A, +)$, donc nécessairement distingué dans A, on sait déjà que l'addition est bien définie, et que le couple $(A/\mathfrak{a}, +)$ est un groupe abélien, de neutre $\overline{0}$. Montrons que la multiplication est bien définie. Supposons que $a, a', b, b' \in A$ vérifient $\overline{a'} = \overline{a}$ et $\overline{b'} = \overline{b}$. Il existe donc $x, y \in \mathfrak{a}$ tels que $a' = a + x, b' = b + y$, et on a alors

$$a'b' = (a + x)(b + y) = ab + ay + xb + xy.$$

Comme x, $y \in \mathfrak{a}$ et que $\mathfrak{a}$ est un idéal bilatère, on a $ay + xb + xy \in \mathfrak{a}$ et donc $\overline{a'b'} = \overline{ab}$, ce qu'il fallait vérifier. Les propriétés de commutativité, associativité, distributivité découlent alors des propriétés correspondantes des lois de A. De plus, $\overline{1}$ est clairement un neutre pour la multiplication. Enfin, le fait que π soit un morphisme d'anneaux provient de la définition même de la multiplication de $A/\mathfrak{a}$. □

1.2. Définition. Soit A un anneau, et soit $\mathfrak{a}$ un idéal bilatère. L'anneau $A/\mathfrak{a}$ est appelé *l'anneau quotient* de A par $\mathfrak{a}$. L'application

$$\pi : A \longrightarrow A/\mathfrak{a}$$

est appelée la *projection canonique*.

On a alors le même genre de propriétés que pour les groupes quotients.

1.3. Proposition. *Soit A un anneau et soit $\mathfrak{a}$ un idéal bilatère. Alors, la projection canonique $\pi : A \longrightarrow A/\mathfrak{a}$ induit une correspondance bijective entre l'ensemble des idéaux à gauche, resp. à droite, resp. bilatères $\mathfrak{a}'$ de A contenant $\mathfrak{a}$ et l'ensemble des idéaux à gauche, resp. à droite, resp. bilatères de $A/\mathfrak{a}$.*

Plus précisément, si $\mathfrak{a}'$ est un idéal à gauche, resp. à droite, resp. bilatère de A contenant $\mathfrak{a}$, alors

$$\mathfrak{a}'/\mathfrak{a} = \pi(\mathfrak{a}') = \{\overline{a'} \mid a' \in \mathfrak{a}'\}$$

est un idéal à gauche, resp. à droite, resp. bilatère de $A/\mathfrak{a}$. Réciproquement, si $\mathfrak{b}$ est un idéal à gauche, resp. à droite, resp. bilatère de $A/\mathfrak{a}$, alors $\pi^{-1}(\mathfrak{b})$ est un idéal à gauche, resp. à droite, resp. bilatère de A contenant $\mathfrak{a}$, et ces deux constructions sont inverses l'une de l'autre.

Démonstration. On fait la démonstration uniquement pour les idéaux à gauche, les deux autres cas se démontrant de manière similaire. D'après la proposition IX-1.6, il reste simplement à voir que $\pi^{-1}(\mathfrak{b})$ est un idéal à gauche de A pour tout idéal à gauche $\mathfrak{b}$ de $A/\mathfrak{a}$, et que $\pi(\mathfrak{a}')$ est un idéal à gauche de $A/\mathfrak{b}$ pour tout idéal à gauche $\mathfrak{a}'$ de A contenant $\mathfrak{a}$.

Puisque π est un morphisme d'anneaux, $\pi^{-1}(\mathfrak{b})$ est un idéal à gauche par le lemme XV-4.5. On sait déjà que $\pi(\mathfrak{a}')$ est un sous-groupe de $A/\mathfrak{a}$ par le lemme VI-2.5 (1). Or, pour tout $\overline{a} \in A/\mathfrak{a}$, et tout $x \in \mathfrak{a}'$, on a

$$\overline{a}\pi(x) = \overline{a}\ \overline{x} = \overline{ax} \in \pi(\mathfrak{a}'),$$

car $x \in \mathfrak{a}'$ et $\mathfrak{a}'$ est un idéal à gauche de A. Ainsi, $\pi(\mathfrak{a}')$ est un idéal à gauche de $A/\mathfrak{a}$. Cela achève la démonstration. □

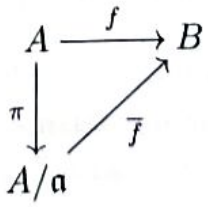
Comme dans le cas des groupes, on a alors à notre disposition un théorème de factorisation, et divers théorèmes d'isomorphisme.

1.4. Théorème (de factorisation). *Soit A un anneau, et soit $\mathfrak{a}$ un idéal bilatère. Soit $f : A \longrightarrow B$ un morphisme d'anneaux tel que $\mathfrak{a} \subset \mathrm{Ker}(f)$. Alors, il existe un unique morphisme d'anneaux*

$$\overline{f} : A/\mathfrak{a} \longrightarrow B$$

tel que $f = \overline{f} \circ \pi$. Ce morphisme d'anneaux est défini par

$$\overline{f}(\overline{x}) = f(x), \quad \text{pour tout } \overline{x} \in A/\mathfrak{a}.$$



De plus, pour tout anneau B, il y a une correspondance bijective naturelle entre l'ensemble $\mathrm{Hom}(A/\mathfrak{a}, B)$ et l'ensemble $\{f \in \mathrm{Hom}(A, B) \mid \mathfrak{a} \subset \mathrm{Ker}(f)\}$.

Elle est donnée par est donnée par

$$\begin{array}{ccc} \varphi & \longmapsto & \varphi \circ \pi \\ \overline{f} & \longmapsfrom & f. \end{array}$$

Démonstration. On sait déjà qu'il existe un unique morphisme de groupes additifs $\overline{f}$ vérifiant les propriétés requises, par le théorème de factorisation pour les groupes, et qu'il est défini par la formule précédente. Il suffit donc vérifier que $\overline{f}$ est aussi un morphisme d'anneaux. Mais on a

$$\overline{f}(\overline{1}_A) = f(1_A) = 1_B,$$

et pour tous $\overline{x}$, $\overline{y} \in A/\mathfrak{a}$, on a

$$\overline{f}(\overline{x})\overline{f}(\overline{y}) = f(x)f(y) = f(xy) = \overline{f}(\overline{xy}) = \overline{f}(\overline{x}\ \overline{y}).$$

D'après le théorème IX-2.1 la correspondance entre $\mathrm{Hom}(A/\mathfrak{a}, B)$ et l'ensemble A' des morphismes $f : A \longrightarrow B$ s'annulant sur $\mathfrak{a}$ est donnée par

$$\begin{array}{ccc} \varphi & \longmapsto & \varphi \circ \pi \\ \overline{f} & \longmapsfrom & f. \end{array}$$

Cela achève la démonstration. □

Soit $f : A \longrightarrow B$ un morphisme d'anneaux. Puisque $\mathrm{Ker}(f)$ est un idéal bilatère de A, on a en particulier un morphisme d'anneaux

$$\overline{f} : A/\mathrm{Ker}(f) \longrightarrow B.$$

Le lemme IX-2.2 entraîne alors le résultat suivant.

1.5. Lemme. *Soit $f : A \longrightarrow B$ un morphisme d'anneaux. Alors, le morphisme d'anneaux*

$$\overline{f} : A/\mathrm{Ker}(f) \longrightarrow B$$

est injectif.

On a alors un théorème d'isomorphisme semblable au théorème IX-2.3.

1.6. Théorème (Premier théorème d'isomorphisme). *Pour tout morphisme d'anneaux $f : A \longrightarrow B$, le morphisme d'anneaux*

$$\overline{f} : A/\mathrm{Ker}(f) \longrightarrow B$$

induit un isomorphisme d'anneaux

$$A/\mathrm{Ker}(f) \simeq \mathrm{Im}(f).$$

Démonstration. D'après le théorème IX-2.3, on sait que $\overline{f}$ est un isomorphisme de groupes, donc une application bijective. Par le théorème de fac-

torisation précédent, on sait aussi que $\overline{f}$ est un morphisme d'anneaux. C'est donc un isomorphisme d'anneaux, d'où le résultat. □

On a également un autre théorème d'isomorphisme, correspondant au troisième théorème d'isomorphisme pour les groupes (le deuxième ne donnant rien au niveau de la structure d'anneau).

1.7. Théorème (Second théorème d'isomorphisme). *Soit A un anneau, et soient $\mathfrak{a}$, $\mathfrak{b}$ deux idéaux bilatères de A. On suppose que $\mathfrak{a} \subset \mathfrak{b}$. Alors, le groupe quotient $\mathfrak{b}/\mathfrak{a}$ est un idéal bilatère de $A/\mathfrak{a}$, et l'on a un isomorphisme d'anneaux*

$$(A/\mathfrak{a})/(\mathfrak{b}/\mathfrak{a}) \simeq A/\mathfrak{b}.$$

Démonstration. La première partie découle de la proposition 1.3. Pour le reste, on procède comme dans la démonstration du théorème IX-2.6 (mais en relevant cette fois que l'on travaille avec des morphismes d'anneaux). □

2. Idéaux premiers, maximaux

Dans ce paragraphe, tous les anneaux seront supposés commutatifs. On parlera donc simplement d'idéal de l'anneau A.

2.1. Définition. Soit A un anneau commutatif. Un idéal $\mathfrak{p}$ de A est dit *premier* si $\mathfrak{p} \neq A$, et si pour tous $a, b \in A$, on a

$$ab \in \mathfrak{p} \implies a \in \mathfrak{p} \text{ ou } b \in \mathfrak{p}.$$

Autrement dit, $\mathfrak{p}$ est premier si $\mathfrak{p} \neq A$, et pour tous éléments $a, b \in A \setminus \mathfrak{p}$, on a $ab \in A \setminus \mathfrak{p}$.

Un idéal $\mathfrak{m}$ est dit *maximal* si $\mathfrak{m} \neq A$ et si pour tout idéal $\mathfrak{a}$ de A, on a

$$\mathfrak{m} \subset \mathfrak{a} \subset A \implies \mathfrak{a} = \mathfrak{m} \text{ ou } \mathfrak{a} = A.$$

Autrement dit, $\mathfrak{m}$ est maximal si c'est un élément maximal de l'ensemble des idéaux de A qui sont distincts de A.

2.2. Exemples

(1) L'idéal (0) est premier si, et seulement si, A est intègre.
Cela provient immédiatement de la définition d'un anneau intègre. Les vérifications sont laissées au lecteur.

(2) Si p est premier, $p\mathbb{Z}$ est un idéal premier de $\mathbb{Z}$. Là encore, on laisse le lecteur vérifier ce fait à titre d'exercice.
L'idéal $p\mathbb{Z}$ est aussi maximal. En effet, d'abord $p\mathbb{Z} \neq \mathbb{Z}$, car 1 n'est pas divisible par p. Supposons ensuite qu'un idéal $\mathfrak{a} = n\mathbb{Z}$, $n \geqslant 0$

vérifie $p\mathbb{Z} \subset n\mathbb{Z} \subset \mathbb{Z}$. Alors, $n|p$, et donc soit $n = 1$ ou p. Ainsi, on a soit $n\mathbb{Z} = \mathbb{Z}$, soit $n\mathbb{Z} = p\mathbb{Z}$.

On veut maintenant démontrer l'existence d'idéaux maximaux, existence qui n'est a priori pas garantie.

2.3. Théorème (Krull). *Soit A un anneau commutatif. Alors, tout idéal $\mathfrak{a} \neq A$ est contenu dans un idéal maximal. En particulier, tout anneau commutatif non trivial possède au moins un idéal maximal.*

Démonstration. Considérons l'ensemble E défini par

$$E = \{\mathfrak{b} \mid \mathfrak{b} \neq A,\ \mathfrak{b} \text{ idéal de } A,\ \mathfrak{a} \subset \mathfrak{b}\}.$$

Alors, E est non vide (car il contient $\mathfrak{a}$), partiellement ordonné par l'inclusion. Soit $(I, \leqslant)$ un ensemble totalement ordonné, et soit $F = (\mathfrak{b}_i)_{i\in I}$ une famille d'éléments de E totalement ordonnée indexée par I. Autrement dit, pour tous $i, j \in I$ tels que $i \leqslant j$, on a $\mathfrak{b}_i \subset \mathfrak{b}_j$. Posons

$$\mathfrak{b} = \bigcup_{i\in I} \mathfrak{b}_i.$$

Alors, $\mathfrak{a} \subset \mathfrak{b}$. De plus, $\mathfrak{b}$ est un idéal de l'anneau A. En effet, si $x, y \in \mathfrak{b}$, il existe des indices $i, j \in I$ tels que $x \in \mathfrak{b}_i$ et $y \in \mathfrak{b}_j$. Soit $k = \max(i, j)$. On a ainsi $i \leqslant k$ et $j \leqslant k$. Puisque F est totalement ordonné, on a alors

$$x \in \mathfrak{b}_i \subset \mathfrak{b}_k \quad \text{et} \quad y \in \mathfrak{b}_j \subset \mathfrak{b}_k.$$

Puisque $\mathfrak{b}_k$ est un idéal, alors pour tout $a \in A$, on a $x + ay \in \mathfrak{b}_k \subset \mathfrak{b}$. Donc $\mathfrak{b}$ est un idéal de A. De plus, on a $\mathfrak{b} \neq A$. Sinon, $\mathfrak{b}$ contiendrait 1. Mais alors, il existerait $i \in I$ tel que $1 \in \mathfrak{b}_i$ et donc on aurait $\mathfrak{b}_i = A$, ce qui contredirait le fait que $\mathfrak{b}_i \in E$. Ainsi $\mathfrak{b} \in E$, et puisque $\mathfrak{b}_i \subset \mathfrak{b}$ pour tout $i \in I$, $\mathfrak{b}$ est un majorant de F. Ainsi, E est inductif, non vide et donc contient un élément maximal $\mathfrak{m}$ par le lemme de Zorn (cf. théorème I-4.6). On a donc $\mathfrak{m} \neq A$ et $\mathfrak{a} \subset \mathfrak{m}$. Montrons que $\mathfrak{m}$ est un idéal maximal. Soit $\mathfrak{b}'$ un idéal de A tel que $\mathfrak{m} \subset \mathfrak{b}' \subset A$. Supposons que $\mathfrak{b}' \neq A$. Alors, on a $\mathfrak{b}' \in E$ et donc $\mathfrak{b}' = \mathfrak{m}$ par maximalité de $\mathfrak{m}$. Cela achève la démonstration du premier point. Pour montrer le dernier point, il suffit de prendre $\mathfrak{a} = (0)$, qui est distinct de A, car A est non trivial. □

Nous allons maintenant établir le lien entre idéaux premiers et éléments nilpotents. On commence par un lemme technique, dont la démonstration repose sur les mêmes idées que celle du théorème de Krull.

2.4. Lemme. *Soit A un anneau commutatif, et soit S une partie non vide de A stable par multiplication. Soit $\mathfrak{a}$ un idéal de A disjoint de S. Alors, $\mathfrak{a}$ est contenu dans un idéal premier disjoint de S.*

En particulier, pour toute partie S non vide stable par multiplication et ne contenant pas 0, il existe un idéal premier de A disjoint de S.

Démonstration. Soit E l'ensemble des idéaux de A contenant $\mathfrak{a}$ et disjoints de S. Alors, E est partiellement ordonné par l'inclusion, et non vide, puisqu'il contient $\mathfrak{a}$. Montrons que E est inductif. Soit $(I, \leqslant)$ un ensemble totalement ordonné, et soit $F = (\mathfrak{a}_i)_{i\in I}$ une famille d'éléments de E totalement ordonnée. Posons

$$\mathfrak{a}' = \bigcup_{i\in I} \mathfrak{a}_i.$$

On a déjà vu dans la démonstration du théorème de Krull que $\mathfrak{a}'$ est un idéal de A. De plus, chaque $\mathfrak{a}_i$ étant disjoint de S et contenant $\mathfrak{a}$, il en est de même de $\mathfrak{a}'$. Comme $\mathfrak{a}_i \subset \mathfrak{a}'$ pour tout $i \in I$, $\mathfrak{a}'$ est donc un majorant de F. Ainsi, E est inductif, non vide, donc admet un élément maximal $\mathfrak{p}$.

Remarquons que $\mathfrak{p}$ contient $\mathfrak{a}$. De plus, $\mathfrak{p} \neq A$, car $A \cap S = S$ est non vide. Supposons que $\mathfrak{p}$ ne soit pas un idéal premier. Il existe donc a, $b \in A \setminus \mathfrak{p}$ tels que $ab \in \mathfrak{p}$. L'idéal $(\mathfrak{p}, a)$ contient strictement $\mathfrak{p}$. Par maximalité de $\mathfrak{p}$, on a donc $(\mathfrak{p}, a) \cap S \neq \emptyset$. Il existe alors $s \in S$, $x \in \mathfrak{p}$ et $y \in A$ tels que $s = x + ay$. De même, il existe $s' \in S$, $x' \in \mathfrak{a}$ et $y' \in A$ tels que $s' = x' + by'$. On a ainsi

$$ss' = xx' + xby' + ayx' + abyy' \in \mathfrak{p},$$

puisque x, x', $ab \in \mathfrak{p}$. Mais alors, $ss' \notin S$, ce qui contredit la stabilité de S par produit.

Le dernier point s'obtient alors en appliquant ce qui précède à l'idéal nul. Cela achève la démonstration. □

On peut maintenant démontrer le résultat suivant.

2.5. Théorème. *Soit A un anneau commutatif. Alors, l'idéal* $\mathrm{Nil}(A)$ *des éléments nilpotents est l'intersection des idéaux premiers de A.*

Démonstration. Soit $x \in \mathrm{Nil}(A)$ et soit $\mathfrak{p}$ un idéal premier. Par définition, il existe $n \geqslant 1$ tel que $x^n = 0$. En particulier, $x^n \in \mathfrak{p}$. Si $x \notin \mathfrak{p}$, la primalité de $\mathfrak{p}$ implique par récurrence que $x^m \notin \mathfrak{p}$ pour tout $m \geqslant 1$. En particulier, on a $0 = x^n \notin \mathfrak{p}$, d'où une contradiction. Ainsi, $x \in \mathfrak{p}$, et $\mathrm{Nil}(A)$ est contenu dans l'intersection des idéaux premiers de A.

Soit maintenant $s \in A \setminus \mathrm{Nil}(A)$. Nous allons montrer qu'il existe un idéal premier $\mathfrak{p}$ ne contenant pas s, ce qui montrera l'inclusion manquante. Soit $S = \{s^n \mid n \geqslant 0\}$. Remarquons que S est stable par multiplication, et

que $0 \notin S$, puisque s n'est pas nilpotent. Par le lemme précédent, il existe un idéal premier $\mathfrak{p}$ disjoint de S. Comme $s \in S$, on a en particulier $s \notin \mathfrak{p}$. Cela achève la démonstration. □

On continue en caractérisant les éléments de A qui appartiennent à au moins un idéal maximal.

2.6. Lemme. *Soit A un anneau commutatif. Alors, $A \setminus A^\times$ est la réunion des idéaux maximaux de A.*

Démonstration. Soit $a \in A$. D'après le lemme XV-4.14 (2), on a $a \in A \setminus A^\times$ si, et seulement si, $(a) \neq A$. D'après le théorème de Krull, ce qui équivaut au fait que (a) est contenu dans un idéal maximal $\mathfrak{m}$. En effet, tout idéal distinct de A est contenu dans un idéal maximal, et A n'est contenu dans aucun idéal maximal par définition. De plus, on a $(a) \subset \mathfrak{m}$ si, et seulement si, $a \in \mathfrak{m}$. Bref, on a montré l'égalité

$$A \setminus A^\times = \bigcup_{\mathfrak{m}} \mathfrak{m},$$

la réunion portant sur l'ensemble des idéaux maximaux. Cela achève la démonstration. □

On achève ce paragraphe en donnant une caractérisation des idéaux premiers et maximaux en termes de quotients.

2.7. Théorème. *Soit A un anneau commutatif.*

(1) *Un idéal $\mathfrak{p}$ est premier si, et seulement si, $A/\mathfrak{p}$ est intègre.*

(2) *Un idéal $\mathfrak{m}$ est maximal si, et seulement si, $A/\mathfrak{m}$ est un corps.*

En particulier, tout idéal maximal est premier.

Démonstration

(1) Supposons que $A/\mathfrak{p}$ soit intègre. En particulier, ce n'est pas l'anneau trivial, et donc $\mathfrak{p} \neq A$. Soient maintenant a, $b \in A$ tels que $ab \in \mathfrak{p}$. Alors, on a

$$\overline{a}\,\overline{b} = \overline{ab} = \overline{0}.$$

Puisque $A/\mathfrak{p}$ est intègre, on a $\overline{a} = \overline{0}$ ou $\overline{b} = \overline{0}$, c'est-à-dire $a \in \mathfrak{p}$ ou $b \in \mathfrak{p}$. Ainsi, $\mathfrak{p}$ est premier.

Inversement, si $\mathfrak{p}$ est premier, alors $\mathfrak{p} \neq A$, donc $A/\mathfrak{p}$ n'est pas trivial. Si maintenant $\overline{a}\,\overline{b} = \overline{0}$ dans $A/\mathfrak{p}$, alors $\overline{ab} = \overline{0}$, et donc $ab \in \mathfrak{p}$. Comme $\mathfrak{p}$ est premier, on obtient $a \in \mathfrak{p}$ ou $b \in \mathfrak{p}$, c'est-à-dire $\overline{a} = \overline{0}$ ou $\overline{b} = \overline{0}$. Ainsi, $A/\mathfrak{p}$ est intègre.

(2) Supposons que $A/\mathfrak{m}$ soit un corps. Encore une fois, l'anneau quotient n'est pas trivial, et $\mathfrak{m} \neq A$. Soit $\mathfrak{a}$ un idéal de A tel que $\mathfrak{m} \subset \mathfrak{a} \subset A$, et supposons que $\mathfrak{a} \neq \mathfrak{m}$. On doit montrer que $\mathfrak{a} = A$. Soit $a \in \mathfrak{a} \setminus \mathfrak{m}$. Alors, la classe $\overline{a}$ est non nulle, et $\overline{a}$ est donc inversible, puisque $A/\mathfrak{m}$ est un corps. Ainsi, il existe une classe $\overline{b} \in A/\mathfrak{m}$ tel que $\overline{a}\,\overline{b} = \overline{1}$. On a donc $\overline{1 - ab} = \overline{0}$, et ainsi $1 - ab \in \mathfrak{m} \subset \mathfrak{a}$. Puisque $a \in \mathfrak{a}$ et $\mathfrak{a}$ est un idéal, on a $ab \in \mathfrak{a}$. Par conséquent, $1 = (1 - ab) + ab \in \mathfrak{a}$. Ainsi, $\mathfrak{a}$ contient un élément inversible, et donc $\mathfrak{a} = A$ par le lemme XV-4.14. Il s'ensuit que $\mathfrak{m}$ est maximal.

Inversement, si $\mathfrak{m}$ est un idéal maximal, alors $\mathfrak{m} \neq A$, et $A/\mathfrak{m}$ est non trivial. Soit $\overline{a} \in (A/\mathfrak{m}) \setminus \{\overline{0}\}$. Ainsi, $a \notin \mathfrak{m}$. Soit $\mathfrak{a} = aA + \mathfrak{m}$; c'est un idéal de A (en fait, c'est l'idéal engendré par a et les éléments de $\mathfrak{m}$). Alors, $\mathfrak{m} \subset \mathfrak{a}$ et $\mathfrak{m} \neq \mathfrak{a}$, puisque $a \in \mathfrak{a} \setminus \mathfrak{m}$. Par maximalité de $\mathfrak{m}$, on obtient $\mathfrak{a} = A$. En particulier, $\mathfrak{a}$ contient 1, et donc il existe $b \in A$ et $x \in \mathfrak{m}$ tels que $1 = ba + x$. Ainsi, $\overline{1} = \overline{b}\overline{a}$ et $\overline{a}$ est inversible. L'anneau $A/\mathfrak{m}$ est donc bien un corps.

Un corps étant intègre par la proposition XV-3.8, on obtient le dernier point. Cela achève la démonstration. □

2.8. Exemples

(1) Les idéaux premiers de $\mathbb{Z}$ sont exactement (0) et les $p\mathbb{Z}$, avec p premier.

(2) Les idéaux maximaux de $\mathbb{Z}$ sont les $p\mathbb{Z}$, avec p premier.

3. Théorème chinois

Nous allons maintenant démontrer un théorème qui permet parfois de décortiquer un peu la structure d'un anneau quotient. On commence par une définition.

3.1. Définition. Soit A un anneau commutatif. On dit que deux idéaux $\mathfrak{a}$ et $\mathfrak{b}$ sont *comaximaux* si l'on a $\mathfrak{a} + \mathfrak{b} = A$.

Cela revient à dire qu'il existe $a \in \mathfrak{a}$ et $b \in \mathfrak{b}$ tels que $a + b = 1$.

3.2. Exemple. Deux idéaux $n\mathbb{Z}$ et $m\mathbb{Z}$ de $\mathbb{Z}$ sont comaximaux si, et seulement si, les entiers n et m sont premiers entre eux, puisque d'après la proposition VI-3.10, on a

$$n\mathbb{Z} + m\mathbb{Z} = \mathrm{pgcd}(n, m)\mathbb{Z}.$$

3.3. Lemme. *Soient $\mathfrak{a}_1, \ldots, \mathfrak{a}_{n+1}$ des idéaux de A deux à deux comaximaux. Alors, $\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n$ et $\mathfrak{a}_{n+1}$ sont comaximaux.*

Démonstration. Puisque $\mathfrak{a}_i + \mathfrak{a}_{n+1} = A$ pour tout $i \in [\![1, n]\!]$ par hypothèse, il existe $a_i \in \mathfrak{a}_i$ et $b_1, \ldots, b_n \in \mathfrak{a}_{n+1}$ tels que

$$a_i + b_i = 1, \quad \text{pour tout } i \in [\![1, n]\!].$$

On a donc

$$(a_1 + b_1) \cdots (a_n + b_n) = 1.$$

En développant et en utilisant le fait que $\mathfrak{a}_{n+1}$ est un idéal, on peut donc écrire

$$a_1 \cdots a_n + b = 1, \quad b \in \mathfrak{a}_{n+1}.$$

Mais, $a_1 \cdots a_n \in \mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n$, car $a_i \in \mathfrak{a}_i$ et chaque $\mathfrak{a}_i$ est un idéal de A, d'où le résultat. □

Notation. Soit $\mathfrak{a}$ un idéal de A, et soient $x, y \in A$. Si $x - y \in \mathfrak{a}$, on notera

$$x \equiv y \mod \mathfrak{a}.$$

Si de plus $\mathfrak{a} = (a)$, on notera aussi

$$x \equiv y \ [a].$$

3.4. Proposition. *Soit A un anneau commutatif, et soient $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ des idéaux de A deux à deux comaximaux. Alors, pour tous $x_1, \ldots, x_n \in A$, il existe $x \in A$ tel que*

$$\begin{cases} x \equiv x_1 \mod \mathfrak{a}_1 \\ \vdots \\ x \equiv x_n \mod \mathfrak{a}_n. \end{cases}$$

Démonstration. On procède par récurrence sur n. Si $n = 1$, x_1 est une solution du système. Supposons que $n = 2$. Puisque $\mathfrak{a}_1$ et $\mathfrak{a}_2$ sont comaximaux, il existe $a_1 \in \mathfrak{a}_1$, $a_2 \in \mathfrak{a}_2$ tels que $a_1 + a_2 = 1$. Soient alors x_1, $x_2 \in A$, et posons $x = x_1 a_2 + x_2 a_1$. Alors, on a

$$x = x_1(1 - a_1) + x_2 a_1 = x_1 + a_1(x_2 - x_1),$$

et donc $x \equiv x_1 \mod \mathfrak{a}_1$. De même, on a

$$x = x_1 a_2 + x_2(1 - a_2) = x_2 + a_2(x_1 - x_2),$$

et donc $x \equiv x_2 \mod \mathfrak{a}_2$. Le théorème est donc vrai pour $n = 2$.

Supposons maintenant que le résultat soit vrai pour $n \geqslant 2$, et montrons le au rang $n+1$. Soient $\mathfrak{a}_1, \ldots, \mathfrak{a}_{n+1}$ des idéaux de A deux à deux comaximaux, et soient $x_1, \ldots, x_{n+1} \in A$. Par hypothèse de récurrence, il existe $y \in A$ tel que

$$y \equiv x_i \mod \mathfrak{a}_i, \quad \text{pour tout } i \in [\![1, n]\!].$$

D'après le lemme précédent, les idéaux $\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n$ et $\mathfrak{a}_{n+1}$ sont comaximaux. D'après le cas $n = 2$, il existe $x \in A$ tel que

$$x \equiv y \mod \mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n \quad \text{et} \quad x \equiv x_{n+1} \mod \mathfrak{a}_{n+1}.$$

Mais alors, on a aussi

$$x \equiv y \equiv x_i \mod \mathfrak{a}_i, \quad \text{pour tout } i \in [\![1, n]\!],$$

et le résultat est vrai au rang $n+1$. Cela achève la récurrence et la démonstration. □

On a alors le résultat suivant.

3.5. Théorème (chinois). *Soient $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ des idéaux de A deux à deux comaximaux. Pour tout $i \in [\![1, n]\!]$, et pour tout $a \in A$, on note $[a]_i$ la classe de a modulo $\mathfrak{a}_i$.*
Alors, le morphisme

$$\begin{aligned} f : A &\longrightarrow A/\mathfrak{a}_1 \times \cdots \times A/\mathfrak{a}_n \\ x &\longmapsto ([x]_1, \ldots, [x]_n) \end{aligned}$$

est surjectif, de noyau $\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n$. En particulier, on a un isomorphisme d'anneaux

$$A/\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n \simeq A/\mathfrak{a}_1 \times \cdots \times A/\mathfrak{a}_n.$$

Démonstration. La surjectivité de f est une reformulation de la proposition précédente. Un élément $x \in A$ est dans le noyau de f si, et seulement si, on a $[x]_i = [0]_i$ pour tout $i \in [\![1, n]\!]$, c'est-à-dire si, et seulement si, $x \in \mathfrak{a}_i$ pour tout $i \in [\![1, n]\!]$. Ainsi, $\mathrm{Ker}(f) = \mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n$.

Le dernier point est une simple application du premier théorème d'isomorphisme à f. □

Nous continuons ces considérations en donnant une formulation un peu différente du théorème chinois, mais qui s'avère extrêmement pratique dans certains cas.

On commence par introduire une nouvelle notation.

Notation. Soit $n \geqslant 0$ un entier. Si $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ sont n idéaux de l'anneau A, on note $\mathfrak{a}_1 \cdots \mathfrak{a}_n$ l'idéal engendré par les éléments de la forme

$$a_1 \cdots a_n, \quad a_i \in \mathfrak{a}_i.$$

Lorsque $\mathfrak{a}_1 = \cdots = \mathfrak{a}_n = \mathfrak{a}$, on le note simplement $\mathfrak{a}^n$.

3.6. Remarques

(1) Lorsque $n = 0$, on obtient l'idéal engendré par le produit vide, à savoir 1, et donc $\mathfrak{a}_1 \cdots \mathfrak{a}_n = A$ si $n = 0$.

(2) le produit d'idéaux est associatif : si $\mathfrak{a}$, $\mathfrak{b}$, $\mathfrak{c}$ sont des idéaux de A, on a

$$(\mathfrak{a}\mathfrak{b})\mathfrak{c} = \mathfrak{a}\mathfrak{b}\mathfrak{c} = \mathfrak{a}(\mathfrak{b}\mathfrak{c}).$$

Il suffit pour s'en convaincre d'écrire explicitement les éléments dans chaque cas.

3.7. Exemple. Si $\mathfrak{a}_i = (\alpha_i)$ pour $i \in [\![1, n]\!]$, on vérifie facilement que

$$\mathfrak{a}_1 \cdots \mathfrak{a}_n = (\alpha_1 \cdots \alpha_n).$$

On a alors le lemme suivant.

3.8. Lemme. *Soit $n \geqslant 0$, et soient $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ des idéaux de A. Alors, on a*

$$\mathfrak{a}_1 \cdots \mathfrak{a}_n \subset \mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n.$$

Si de plus $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ sont deux à deux comaximaux, alors on a

$$\mathfrak{a}_1 \cdots \mathfrak{a}_n = \mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n.$$

Démonstration. Montrons la première partie. Pour cela, il suffit de vérifier que pour tous $a_1 \in \mathfrak{a}_1, \ldots, a_n \in \mathfrak{a}_n$, on a $a_1 \cdots a_n \in \mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n$, ce qui est clair, car $\mathfrak{a}_i$ est un idéal. Nous allons démontrer la seconde partie par récurrence. Il reste simplement à démontrer l'inclusion non triviale

$$\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n \subset \mathfrak{a}_1 \cdots \mathfrak{a}_n.$$

Si $n = 0$ ou 1, il n'y a rien à faire. Montrons-le pour $n = 2$. Les idéaux $\mathfrak{a}_1$ et $\mathfrak{a}_2$ étant comaximaux, il existe $a_1 \in \mathfrak{a}_1$ et $a_2 \in \mathfrak{a}_2$ tels que $a_1 + a_2 = 1$.

Soit maintenant $x \in \mathfrak{a}_1 \cap \mathfrak{a}_2$. On a alors

$$x = x\,1 = xa_1 + xa_2 = a_1x + xa_2.$$

Or, $a_1x, xa_2 \in \mathfrak{a}_1\mathfrak{a}_2$ par définition, puisque $x \in \mathfrak{a}_1 \cap \mathfrak{a}_2$. Mais alors, $x \in \mathfrak{a}_1\mathfrak{a}_2$, et l'on a fini.

Supposons maintenant avoir montré le résultat pour un produit de n idéaux, avec $n \geqslant 2$, et soient $\mathfrak{a}_1, \ldots, \mathfrak{a}_{n+1}$, $n+1$ idéaux deux à deux comaximaux. D'après le lemme 3.3, les idéaux $\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n$ et $\mathfrak{a}_{n+1}$ sont comaximaux. En utilisant le cas $n = 2$, puis l'hypothèse de récurrence, on obtient donc

$$\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n \cap \mathfrak{a}_{n+1} = (\mathfrak{a}_1 \cap \cdots \cap \mathfrak{a}_n)\mathfrak{a}_{n+1} = (\mathfrak{a}_1 \cdots \mathfrak{a}_n)\mathfrak{a}_{n+1}.$$

On achève alors la récurrence en utilisant la remarque 3.6 (2). □

On obtient alors une nouvelle formulation du théorème chinois.

3.9. Théorème. *Soit A un anneau, et soient $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ des idéaux de A deux à deux comaximaux. Pour tout $i \in [\![1, n]\!]$, et tout $a \in A$, on note $[a]_i$ la classe de a modulo $\mathfrak{a}_i$.*
Alors, le morphisme

$$\begin{aligned} f\colon A &\longrightarrow A/\mathfrak{a}_1 \times \cdots \times A/\mathfrak{a}_n \\ x &\longmapsto \quad ([x]_1, \ldots, [x]_n) \end{aligned}$$

est surjectif, de noyau $\mathfrak{a}_1 \cdots \mathfrak{a}_n$. En particulier, on a un isomorphisme d'anneaux

$$A/\mathfrak{a}_1 \cdots \mathfrak{a}_n \simeq A/\mathfrak{a}_1 \times \cdots \times A/\mathfrak{a}_n.$$

Lorsque $\mathfrak{a} = (a)$ et $\mathfrak{b} = (b)$, les idéaux $\mathfrak{a}$ et $\mathfrak{b}$ sont comaximaux si, et seulement si, il existe $u, v \in A$ tels que $ua + vb = 1$. De tels éléments méritent un nom particulier.

3.10. Définition. Soit A un anneau commutatif. Deux éléments $a, b \in A$ sont *étrangers* s'il existe $u, v \in A$ tels que $ua + vb = 1$.

3.11. Exemple. Deux entiers $n, m \in \mathbb{Z}$ sont étrangers si, et seulement si, ils sont premiers entre eux, d'après le théorème de Bézout [1].

3.12. Remarque. Deux éléments $a, b \in A$ sont étrangers si, et seulement si, a est inversible modulo b, c'est-à-dire $\overline{a} \in \big(A/(b)\big)^\times$.

En effet, si a est inversible modulo b, il existe $u \in A$ tel que $ua \equiv 1\ [b]$, soit $1 - ua \equiv 0\ [b]$. Mais alors, il existe $v \in A$ tel que $1 - ua = vb$, ce qui se récrit $ua + vb = 1$. Par conséquent, a et b sont étrangers.

Inversement, s'il existe $u, v \in A$ tels que $ua + vb = 1$, alors on a $ua \equiv 1\ [b]$, et a est donc inversible modulo b.

Dans le cas où tous les idéaux $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ sont principaux (par exemple lorsque A est principal), le théorème chinois se retraduit alors comme suit.

3.13. Théorème. *Soit A un anneau commutatif, et soient $a_1, \ldots, a_n \in A$ deux à deux étrangers. Pour tout $i \in [\![1, n]\!]$, et pour tout $a \in A$, on note $[a]_i$ la classe de a modulo $\mathfrak{a}_i$. Alors, le morphisme*

$$\begin{aligned} f\colon A &\longrightarrow A/(a_1) \times \cdots \times A/(a_n) \\ x &\longmapsto \quad ([x]_1, \ldots, [x]_n) \end{aligned}$$

est surjectif, de noyau $(a_1 \cdots a_n)$.

1. On verra néanmoins dans le chapitre suivant que cela n'est pas toujours le cas dans tous les anneaux, une fois les notions de divisibilité et d'éléments premiers entre eux généralisées convenablement.

En particulier, on a un isomorphisme d'anneaux

$$A/(a_1 \cdots a_n) \simeq A/(a_1) \times \cdots \times A/(a_n).$$

Nous allons finir ce chapitre par quelques illustrations calculatoires des théorèmes précédents.

Soient $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$ des idéaux de A deux à deux comaximaux. La démonstration de la proposition 3.4 donne une méthode pour résoudre un système de n congruences

$$\begin{cases} x \equiv x_1 \mod \mathfrak{a}_1 \\ \quad \vdots \\ x \equiv x_n \mod \mathfrak{a}_n. \end{cases}$$

On commence par résoudre le système formé par les deux premières équations

$$\begin{cases} x \equiv x_1 \mod \mathfrak{a}_1 \\ x \equiv x_2 \mod \mathfrak{a}_2, \end{cases}$$

en s'aidant d'une relation de comaximalité, comme expliqué dans la démonstration. On obtient alors une solution y_1, unique modulo $\mathfrak{a}_1 \cap \mathfrak{a}_2$. Le système de départ est alors équivalent au système de $n-1$ congruences

$$\begin{cases} x \equiv y_1 \mod \mathfrak{a}_1 \cap \mathfrak{a}_2 \\ x \equiv x_3 \mod \mathfrak{a}_3 \\ \quad \vdots \\ x \equiv x_n \mod \mathfrak{a}_n. \end{cases}$$

On résout alors le système

$$\begin{cases} x \equiv y_1 \mod \mathfrak{a}_1 \cap \mathfrak{a}_2 \\ x \equiv x_3 \mod \mathfrak{a}_3, \end{cases}$$

en s'appuyant sur la méthode de démonstration du lemme 3.8 pour écrire la relation de comaximalité entre $\mathfrak{a}_1 \cap \mathfrak{a}_2$ et $\mathfrak{a}_3$. On obtient alors de proche en proche une solution du système de départ.

La résolution d'un tel système de congruences est particulièrement simple à mettre en œuvre dans le cas où les idéaux sont tous principaux. Considérons en effet le cas d'un système de deux congruences

$$\begin{cases} x \equiv x_1 \ [a_1] \\ x \equiv x_2 \ [a_2], \end{cases}$$

où a_1 et a_2 sont étrangers.

D'après la première équation, une solution $x \in A$ est de la forme

$$x = x_1 + a_1 y, \quad y \in A.$$

En reportant dans la seconde équation, on obtient

$$a_1 y \equiv x_2 - x_1 \ [a_2].$$

Puisque a_1 et a_2 sont étrangers, il existe donc $u, v \in A$ tels que

$$ua_1 + va_2 = 1.$$

En particulier, $ua_1 \equiv 1 \ [a_2]$. Autrement dit, u est un inverse de a_1 modulo a_2. En multipliant la congruence précédente par u, on obtient

$$y \equiv u(x_2 - x_1) \ [a_2].$$

Ainsi, il existe $z \in A$ tel que $y = u(x_2 - x_1) + za_2$. On obtient alors

$$x = x_1 + a_1 u(x_2 - x_1) + a_1 a_2 z, \quad z \in A.$$

Si l'on a plus de deux congruences, on résout le système de proche en proche.

3.14. Exemple. Soit $A = \mathbb{Z}$. Résolvons le système

$$\begin{cases} x \equiv 1 \ [3] \\ x \equiv 2 \ [4] \\ x \equiv -1 \ [7]. \end{cases}$$

D'après la première équation, on a $x = 1 + 3y, \ y \in \mathbb{Z}$. En reportant dans la deuxième équation, on obtient

$$3y \equiv 1 \ [4].$$

L'inverse de 3 modulo 4 est -1, comme on le voit facilement (il n'y a pas besoin ici d'appliquer l'algorithme d'Euclide pour trouver une relation de Bézout). On a alors

$$y \equiv -1 \ [4],$$

soit $y = -1 + 4z, \ z \in \mathbb{Z}$, et en reportant dans l'égalité ci-dessus, on obtient que $x = 1 + 3(-1 + 4z) = -2 + 12z$, avec $z \in \mathbb{Z}$.

En injectant cela dans la dernière équation, on obtient $12z \equiv 1 \ [7]$, soit

$$5z \equiv 1 \ [7].$$

L'inverse de 5 modulo 7 est 3, et l'on a ainsi

$$z \equiv 3 \ [7],$$

soit $z = 3 + 7t, \ t \in \mathbb{Z}$. On obtient finalement

$$x = -2 + 12(3 + 7t) = 34 + 84t, \quad t \in \mathbb{Z}.$$

3.15. Remarque. La seule difficulté de la méthode de résolution précédente est de savoir inverser un élément a modulo b, lorsque c'est possible. Cela revient à savoir calculer une relation du type $ua + vb = 1$. Dans le cas où A est l'anneau $\mathbb{Z}$, il suffit d'appliquer l'algorithme d'Euclide. Cela marche aussi parfaitement dans le cas où $A = K[X]$ (K corps), comme le sait pertinemment le lecteur [2].

Remarquons au passage qu'avec cette méthode, on n'est pas obligé de reprendre les calculs depuis le début si l'on rajoute une équation au système initial. En revanche, lorsque l'on a plusieurs systèmes de congruences à résoudre, relativement aux mêmes idéaux $\mathfrak{a}_1, \ldots, \mathfrak{a}_n$, il peut être judicieux de calculer l'inverse de l'isomorphisme

$$\varphi : A/(\mathfrak{a}_1 \cdots \mathfrak{a}_n) \xrightarrow{\sim} A/\mathfrak{a}_1 \times \cdots \times A/\mathfrak{a}_n.$$

Cet isomorphisme envoie la classe $\overline{x}$ sur $([x]_1, \ldots, [x]_n)$. En particulier, la composition avec la projection sur la i-ème composante est le morphisme

$$\begin{aligned} A/(\mathfrak{a}_1 \cdots \mathfrak{a}_n) &\longrightarrow A/\mathfrak{a}_i \\ \overline{x} &\longmapsto [x]_i. \end{aligned}$$

D'après la remarque XV-5.6, si l'on pose

$$\overline{e}_i = \varphi^{-1}(\varepsilon_i), \quad \text{avec } \varepsilon_i = \big(\pi_1(0), \ldots, \pi_{i-1}(0), \pi_i(1), \pi_{i+1}(0), \ldots, \pi_n(0)\big),$$

l'isomorphisme inverse de φ est donné par

$$\begin{aligned} \varphi^{-1} : A/\mathfrak{a}_1 \times \cdots \times A/\mathfrak{a}_n &\longrightarrow A/(\mathfrak{a}_1 \cdots \mathfrak{a}_n) \\ ([x_1]_1, \ldots, [x_n]_n) &\longmapsto \sum_{i=1}^{n} \overline{e}_i \overline{x}_i. \end{aligned}$$

On peut d'ailleurs retrouver ce résultat de manière directe en remarquant que l'on a

$$\begin{aligned} \big([x_1]_1, \ldots, [x_n]_n\big) &= \varepsilon_1\big([x_1]_1, \ldots, [x_1]_n\big) + \cdots + \varepsilon_n\big([x_n]_1, \ldots, [x_n]_n\big) \\ &= \varphi(\overline{e}_1)\varphi(\overline{x}_1) + \cdots + \varphi(\overline{e}_n)\varphi(\overline{x}_n) \\ &= \varphi\Big(\sum_{i=1}^{n} \overline{e}_i \overline{x}_i\Big). \end{aligned}$$

Autrement dit, il suffit de trouver pour tout $i \in [\![1, n]\!]$ une solution $e_i \in A$ du système de n congruences

$$x \equiv \delta_{ij} \mod \mathfrak{a}_j, \quad j \in \in [\![1, n]\!]$$

pour calculer explicitement φ^{-1}, et ainsi trouver une solution de n'importe quel système de congruences.

2. Au moins dans le cas où $K = \mathbb{R}$ ou $\mathbb{C}$.

Or, un élément $x \in A$ est solution du système précédent si, et seulement si, on a $x \in \bigcap_{j \neq i} \mathfrak{a}_j$ et $x \equiv 1 \mod \mathfrak{a}_i$. Ainsi, ce système est équivalent au système de deux congruences

$$\begin{cases} x \equiv 0 \mod \bigcap_{j \neq i} \mathfrak{a}_j \\ x \equiv 1 \mod \mathfrak{a}_i, \end{cases}$$

que l'on peut alors résoudre alors comme déjà expliqué auparavant.

Encore une fois, cela est particulièrement simple à mettre en œuvre lorsque tous les idéaux sont principaux. Dans ce cas, le système

$$x \equiv \delta_{ij} \ [a_i], \quad j \in [\![1, n]\!]$$

est équivalent au système

$$\begin{cases} x \equiv 0 \quad \left[\prod_{j \neq i} a_j\right] \\ x \equiv 1 \quad [a_i]\,. \end{cases}$$

3.16. Exemple. Calculons l'inverse de l'isomorphisme

$$\varphi : \mathbb{Z}/84\mathbb{Z} \xrightarrow{\sim} \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/7\mathbb{Z}.$$

Considérons le système de congruences

$$\begin{cases} x \equiv 1 \quad [3] \\ x \equiv 0 \quad [4] \\ x \equiv 0 \quad [7]\,. \end{cases}$$

Les deux dernières congruences signifient que x est divisible par 4 et 7, donc par 28, car 4 et 7 sont premiers entre eux [3]. Ainsi, ce système est équivalent au système de congruences

$$\begin{cases} x \equiv 1 \quad [3] \\ x \equiv 0 \quad [28]. \end{cases}$$

D'après la seconde équation, on a $x = 28y$, avec $y \in \mathbb{Z}$, et en reportant dans la première équation, on a

$$28y \equiv 1 \quad [3],$$

soit $y \equiv 1 \ [3]$. Par conséquent, $e_1 = 28$ est une solution du système de départ, et $\overline{e}_1 = \overline{28}$ est donc le premier idempotent recherché. On calcule de même les deux autres idempotents, et l'on trouve $\overline{e}_2 = \overline{21}$ et $\overline{e}_3 = \overline{36}$.

3. On peut aussi appliquer le lemme 3.8 et son interprétation dans le cas des idéaux principaux, mais il est tellement plus simple de faire ce raisonnement arithmétique !

Si $\hat{a}$, $\dot{a}$, $\tilde{a}$ et $\overline{a}$ désignent les classes de $a \in \mathbb{Z}$ modulo 3, 4, 7 et 84, respectivement, on a ainsi

$$\begin{aligned} \varphi^{-1} \colon \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/4\mathbb{Z} \times \mathbb{Z}/7\mathbb{Z} &\longrightarrow \mathbb{Z}/84\mathbb{Z} \\ (\hat{a}_1, \dot{a}_2, \tilde{a}_3) &\longmapsto \overline{28a_1 + 21a_2 + 36a_3}\,. \end{aligned}$$

En particulier, on obtient

$$\varphi^{-1}\big((\hat{1}, \dot{2}, \widetilde{-1})\big) = \overline{34},$$

ce qui est à la reformulation près le résultat de l'exemple 3.14.

Tout cela fonctionne bien entendu de la même manière dans le cas des polynômes à coefficients dans un corps K.

4. Exercices

E1. Sur un bateau, il y a dix-sept pirates et un cuisinier. Les pirates ont volé un coffre de pièces d'or, dont ils décident de partager les pièces équitablement entre eux. Au terme de ce partage, il reste trois pièces, qu'ils projettent de donner au cuisinier. Mais le bateau essuie alors une tempête, au cours de laquelle six pirates sont noyés. Les pirates survivants recommencent le partage, au terme duquel il doit rester quatre pièces pour le cuisinier. Une effroyable dispute éclate alors entre les pirates, et cinq pirates y perdent la vie. Les pirates restants recommencent encore le partage, qui doit alors laisser cinq pièces au cuisinier. Lassé d'attendre, celui-ci décide d'empoisonner le reste des pirates et de garder le coffre pour lui : combien de pièces peut-il espérer au minimum ?

E2. Trouver l'isomorphisme inverse de l'isomorphisme canonique

$$\mathbb{Z}/30\mathbb{Z} \xrightarrow{\sim} \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/5\mathbb{Z},$$

ainsi que l'isomorphisme inverse de

$$\mathbb{R}[X]/(X^4 + X^2) \xrightarrow{\sim} \mathbb{R}[X]/(X^2) \times \mathbb{R}[X]/(X^2 + 1).$$

E3. Soit A un anneau commutatif, et soit $\mathfrak{a}$ un idéal de A. On pose

$$\sqrt{\mathfrak{a}} = \{x \in A \mid \text{ il existe } n \geqslant 1 \text{ tel que } x^n \in \mathfrak{a}\}.$$

a. Montrer que $\sqrt{\mathfrak{a}}$ est un idéal de A contenant $\mathfrak{a}$, et que

$$\sqrt{\sqrt{\mathfrak{a}}} = \sqrt{\mathfrak{a}}.$$

b. Calculer $\sqrt{\mathfrak{a}}$ pour $A = \mathbb{Z}$ et $\mathfrak{a} = (0)$, $\mathbb{Z}$, $4\mathbb{Z}$, $8\mathbb{Z}$, $9\mathbb{Z}$, $24\mathbb{Z}$, $15\mathbb{Z}$, $m\mathbb{Z}$.

c. Calculer $\sqrt{\mathfrak{a}}$ lorsque $\mathfrak{a} = \mathfrak{p}^m$, où $\mathfrak{p}$ est un idéal premier et $m \geqslant 1$.

d. Montrer que $A/\sqrt{\mathfrak{a}}$ n'a pas d'éléments nilpotents non nuls.

E4. Soit $\omega \in \mathbb{C}$. Déterminer le noyau du morphisme d'évaluation

$$ev_\omega : \mathbb{Q}[X] \longrightarrow \mathbb{C}$$

dans les cas suivants : $\omega = i$, $\omega = \sqrt{2}$, $\omega = j$, $\omega = e^{i\pi/4}$.

Que donne le théorème de factorisation ? Et si l'on remplace $\mathbb{Q}$ par $\mathbb{Z}$?

Indication. Dans chaque cas, on cherchera un polynôme unitaire P_0 à coefficients dans $\mathbb{Q}$ ou $\mathbb{Z}$ qui annule ω, de degré le plus petit possible, et l'on effectuera la division euclidienne de $P \in \mathbb{Q}[X]$ ou $\mathbb{Z}[X]$ par P_0.

E5. Identifier les anneaux quotients suivants à un anneau plus simple.

(1) $A[X]/(X-a)$ $(a \in A)$;

(2) $\mathbb{Z}[X]/(2X-1)$.

E6. Dans les cas suivants, l'idéal $\mathfrak{a}$ est-il premier ? maximal ?

(1) $A = \mathbb{Q}[X]$, $\mathfrak{a} = (X^2-2)$,

(2) $A = \mathbb{Z}[X]$, $\mathfrak{a} = (X^2-2)$,

(3) $A = \mathbb{R}[X]$, $\mathfrak{a} = (X^2-2)$,

(4) $A = \mathbb{R}[X]$, $\mathfrak{a} = (X^2-1)$,

(5) $A = \mathbb{R}[X]$, $\mathfrak{a} = (X^2+1)$,

(6) $A = \mathbb{C}[X]$, $\mathfrak{a} = (X^2+1)$,

(7) $A = \mathbb{Z}[X]$, $\mathfrak{a} = (2, X)$,

(8) $A = \mathbb{Z}[X]$, $\mathfrak{a} = (m, X)$, $m \geqslant 1$ fixé,

(9) $A = \mathbb{R}[X, Y]$, $\mathfrak{a} = (X+Y)$,

(10) $A = \mathbb{R}[X, Y]$, $\mathfrak{a} = (X, Y)$.

E7. Soit A un anneau quelconque, et soit $\mathfrak{m}$ un idéal bilatère maximal pour l'inclusion parmi les idéaux bilatères $\neq A$. L'anneau quotient $A/\mathfrak{m}$ est-il un anneau à division ?

Indication. Utiliser l'exercice XV-30.

E8. On suppose connus les définitions et résultats de l'exercice XV-13. Soit A un anneau de Boole (donc commutatif).

a. Soit $\mathfrak{a}$ un idéal de A. Montrer que $A/\mathfrak{a}$ est un anneau de Boole.

b. Soit $\mathfrak{p}$ un idéal premier de A. Montrer que $A/\mathfrak{p} \simeq \mathbb{Z}/2\mathbb{Z}$. En déduire que tout idéal premier d'un anneau de Boole est maximal, et que pour tous $x, y \in A \setminus \mathfrak{p}$, on a $x + y \in \mathfrak{p}$.

c. On note $\mathrm{Spec}(A)$ l'ensemble des idéaux premiers de A. Pour tout $a \in A$, on pose

$$D(a) = \{\mathfrak{p} \in \mathrm{Spec}(A) \mid a \notin \mathfrak{p}\}.$$

(i) Montrer que $D(1) = \mathrm{Spec}(A)$, $D(0) = \emptyset$, et que pour tous $a, a' \in A$, on a

$$D(a+a') = D(a)\Delta D(a') \quad \text{et} \quad D(aa') = D(a) \cap D(a').$$

(ii) Soit $a \in A \setminus \{0\}$. Justifier que $1-a$ n'est pas inversible. En déduire qu'il existe $\mathfrak{p} \in \mathrm{Spec}(A)$ tel que $1-a \in \mathfrak{p}$, puis que $D(a) \neq \emptyset$.

(iii) En déduire que l'application $D : A \longrightarrow \mathscr{P}\big(\mathrm{Spec}(A)\big)$ est un morphisme d'anneaux injectif, et donc que tout anneau de Boole est isomorphe à un sous-anneau de $\mathscr{P}(E)$ pour un certain ensemble E.

d. On suppose que A est un anneau de Boole fini.

(i) Pour tout $\mathfrak{p} \in \mathrm{Spec}(A)$, montrer que $D\Big(\prod_{a \notin \mathfrak{p}} a\Big) = \mathfrak{p}$.

(ii) En déduire que $A \simeq \mathscr{P}\big(\mathrm{Spec}(A)\big)$.

e. Montrer que tout anneau de Boole fini est isomorphe à $(\mathbb{Z}/2\mathbb{Z})^n$ pour un certain $n \geqslant 0$.

E9. Vrai ou faux ?

a. Soit B un anneau, et soit A un sous-anneau de B.

(i) Si $\mathfrak{P}$ est un idéal premier de B, $\mathfrak{P} \cap A$ est un idéal premier de A.

(ii) Si $\mathfrak{M}$ est un idéal maximal de B, $\mathfrak{M} \cap A$ est un idéal maximal de A.

b. Soit $f : A \longrightarrow B$ un morphisme d'anneaux.

(i) Si $\mathfrak{P}$ est un idéal premier de B, $f^{-1}(\mathfrak{P})$ est un idéal premier de A.

(ii) Si $\mathfrak{M}$ est un idéal maximal de B, $f^{-1}(\mathfrak{M})$ est un idéal maximal de A.

E10. Soit K un corps, et soit $P \in K[X]$ un polynôme de degré $d \geqslant 1$.

a. Montrer que la loi externe

$$\begin{array}{ccc} K \times K[X]/(P) & \longrightarrow & K[X]/(P) \\ (\lambda, \overline{Q}) & \longmapsto & \overline{\lambda Q} \end{array}$$

est bien définie, et induit une structure de K-espace vectoriel sur l'anneau $K[X]/(P)$. Montrer que $\{\overline{1}, \overline{X}, \ldots, \overline{X^{d-1}}\}$ en est une base.

b. Justifier que le morphisme $K \longrightarrow K[X]/(P)$ est injectif. On identifie K à son image dans $K[X]/(P)$. Montrer que le polynôme P, vu comme polynôme à coefficients dans $K[X]/(P)$, a une racine dans l'anneau $K[X]/(P)$.

c. Montrer que pour tout $\lambda \in K$, on a

$$K[X]/\big(P(X-\lambda)\big) \simeq K[X]/(P).$$

Est-ce encore vrai si l'on remplace K par un anneau commutatif A ?

E11. Soit K un corps tel que $\mathrm{car}(K) \neq 2$, et soit $P = X^2 + aX + b \in K[X]$. On pose $\Delta = a^2 - 4b$.

a. Montrer que l'on a un isomorphisme d'anneaux

$$K[X]/(P) \simeq K[X]/(X^2 - \Delta).$$

b. Si $\Delta = 0$, montrer que P possède une racine double dans K et que l'on a $K[X]/(P) \simeq K[X]/(X^2)$.

c. Si $\Delta \in K^{\times 2} = \{\lambda^2 \mid \lambda \in K^\times\}$, montrer que P possède deux racines distinctes dans K et que l'on a

$$K[X]/(P) \simeq K \times K.$$

d. Si $\Delta \in K^\times \setminus K^{\times 2}$, montrer que P n'a aucune racine dans K et que l'anneau $K[X]/(P)$ est un corps.

Indication. Si $\Delta \in K^\times \setminus K^{\times 2}$, montrer que $K[X]/(X^2 - \Delta)$ est un corps.

E12. Le but de cet exercice est de classifier les anneaux à p^2 éléments à isomorphisme près, où p est un nombre premier. On pourra utiliser les résultats des exercices 10 et 11.

a. Soit $P \in \mathbb{F}_p[X]$ un polynôme unitaire de degré 2. Montrer que $\mathbb{F}_p[X]/(P)$ est un anneau à p^2 éléments, dont tous les éléments non nuls sont d'ordre p. Peut-il être isomorphe à $\mathbb{Z}/p^2\mathbb{Z}$?

b. On suppose $p = 2$. Pour chaque polynôme P, déterminer si $\mathbb{F}_2[X]/(P)$ est intègre ou pas, un corps ou pas.

Dans la suite, A désigne un anneau commutatif à p^2 éléments.

c. Justifier que $\mathrm{car}(A) = p$ ou p^2.

d. On suppose que $\mathrm{car}(A) = p^2$. Montrer que $A \simeq \mathbb{Z}/p^2\mathbb{Z}$.

e. On suppose que $\mathrm{car}(A) = p$.

(i) Justifier l'existence de $a \in A \setminus \langle 1_A \rangle$. Vérifier que l'on a

$$A = \{m_1 \cdot 1_A + m_2 \cdot a \mid m_1,\, m_2 \in [\![0, p-1]\!]\}.$$

Indication. Comparer le nombre d'éléments des deux ensembles.

(ii) Justifier l'existence d'un morphisme surjectif $\varphi : \mathbb{F}_p[X] \longrightarrow A$ tel que $\varphi(\overline{m}) = m \cdot 1_A$ pour tout $\overline{m} \in \mathbb{F}_p$ et $\varphi(X) = a$.

(iii) En considérant a^2, justifier l'existence d'un polynôme $P \in \mathbb{F}_p[X]$ unitaire de degré 2 tel que $\varphi(P) = 0$. Montrer alors que l'on a

$$A \simeq \mathbb{F}_p[X]/(P).$$

f. Montrer qu'un anneau commutatif à quatre éléments est isomorphe à un des anneaux suivants :

$$\mathbb{Z}/4\mathbb{Z}, \quad \mathbb{F}_2 \times \mathbb{F}_2, \quad \mathbb{F}_2[X]/(X^2), \quad \mathbb{F}_2[X]/(X^2 + X + \overline{1}).$$

Montrer que tous ces anneaux sont deux à deux non isomorphes. Combien y a-t-il de corps à 4 éléments à isomorphisme près ? Comparer les tables de multiplication des anneaux précédents à celles trouvées dans l'exercice XV-10.

g. Soit p un nombre premier impair.

On admet que le groupe quotient $\mathbb{F}_p^\times/\mathbb{F}_p^{\times 2}$ possède deux éléments (cf. lemme XVII-4.1), et l'on fixe un élément $\overline{\varepsilon} \in \mathbb{F}_p^\times \setminus \mathbb{F}_p^{\times 2}$. Montrer qu'un anneau commutatif à p^2 éléments est isomorphe à un des anneaux suivants :

$$\mathbb{Z}/p^2\mathbb{Z}, \quad \mathbb{F}_p \times \mathbb{F}_p, \quad \mathbb{F}_p[X]/(X^2), \quad \mathbb{F}_p[X]/(X^2 - \overline{\varepsilon}).$$

Montrer que ces anneaux sont deux à deux non isomorphes. Combien y a-t-il de corps à p^2 éléments à isomorphisme près ?

E13. Soit A un anneau commutatif non trivial. On suppose que A n'est pas isomorphe à $\mathbb{Z}/4\mathbb{Z}$ ou à $\mathbb{F}_2[X]/(X^2)$. Montrer que les propriétés suivantes sont équivalentes :

(1) l'anneau A est intègre ;

(2) tout polynôme unitaire de $A[X]$ de degré n a au plus n racines dans A ;

(3) tout polynôme unitaire de $A[X]$ de degré 2 a au plus 2 racines dans A.

Indication. Soit A un anneau non intègre vérifiant (3), et soit $a \in A\setminus\{0\}$ un diviseur de zéro. En considérant le polynôme $X(X-a+x)$, montrer que pour tout $x \in A$ tel que $ax = 0$, on a $x = 0$ ou $x = a$. En déduire que pour tout $a' \in A \setminus \{0\}$, on a $a'a = a$ ou $a'a = 0$. Montrer enfin que pour tout $a' \in A \setminus \{0, a\}$, on a $a' = 1$ ou $1 + a$, et conclure que A est isomorphe à $\mathbb{Z}/4\mathbb{Z}$ et $\mathbb{F}_2[X]/(X^2)$.

E14. Soit $A = \mathbb{Z}[i]/(2+i)$.

a. Soit $m \in \mathbb{Z}$. Trouver une condition nécessaire et suffisante sur m pour que l'équation $(2+i)z = m$ ait au moins une solution $z \in \mathbb{Z}[i]$.

b. En déduire la caractéristique de A.

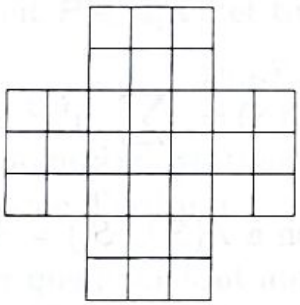
c. Montrer que le morphisme $\Theta_A : \mathbb{Z} \longrightarrow A$ est surjectif.

d. L'idéal $(2+i)$ est-il premier ? maximal ?

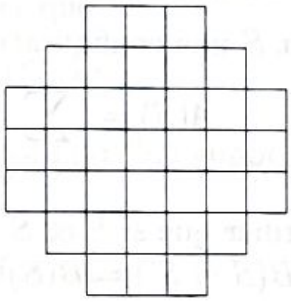
e. Reprendre les questions précédentes avec $A = \mathbb{Z}[j]/(2-5j)$.

E15. On considère un jeu de solitaire, c'est-à-dire un plateau d'un seul tenant formé de cases carrées. Par exemple,

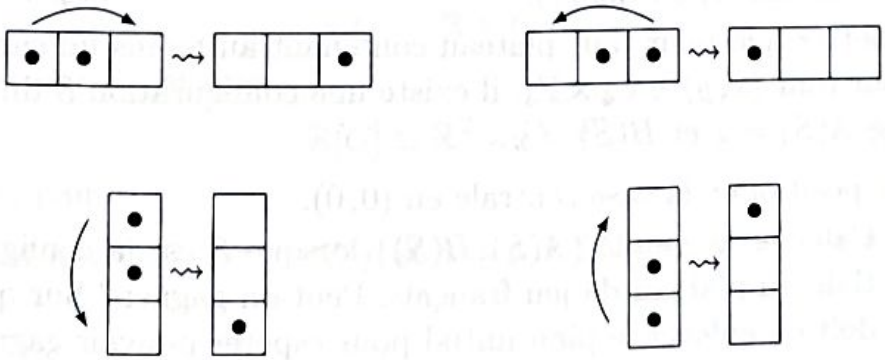
le plateau anglais



le plateau français



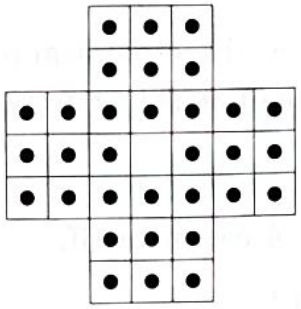
Une configuration S est un ensemble de cases du plateau contenant un pion. On déplace un pion sur une case vide en sautant par dessus un pion (dans les directions N, S, O ou E). On enlève alors le pion par dessus duquel on a sauté :



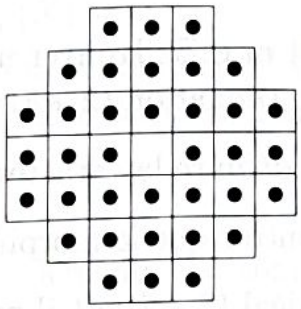
Le but du solitaire est d'enlever tous les pions sauf un, en partant d'une configuration initiale donnée (contenant au moins une case vide). Une partie est gagnée lorsqu'il reste un seul pion sur le plateau après un nombre fini de coups.

Par exemple, la configuration initiale du jeu de solitaire est celle donnée ci-après.

jeu anglais



jeu français



On peut étiqueter les cases du plateau en considérant que les pions sont placés sur des points de $\mathbb{Z}^2$, et ainsi identifier toute configuration à un sous-ensemble de $\mathbb{Z}^2$. On rappelle que $\mathbb{F}_4 = \mathbb{F}_2[X]/(X^2+X+\overline{1})$ est un corps à quatre éléments. On note α la classe de X modulo $X^2+X+\overline{1}$.

a. Donner les éléments de $\mathbb{F}_4$, ainsi que les tables d'addition et de multiplication.

Soit S une configuration. On pose

$$A(S) = \sum_{(i,j)\in S} \alpha^{i+j} \in \mathbb{F}_4, \qquad B(S) = \sum_{(i,j)\in S} \alpha^{i-j} \in \mathbb{F}_4.$$

b. Vérifier que si S et S' sont disjoints, on a $A(S \cup S') = A(S) + A(S')$ et $B(S \cup S') = B(S) + B(S')$.

c. Montrer que si l'on peut passer de la configuration S à la configuration T par une succession de mouvements autorisés, alors on a

$$A(S) = A(T) \quad \text{et} \quad B(S) = B(T).$$

Indication. Il suffit de traiter le cas où l'on passe de S à T en faisant un seul saut (pourquoi ?).

d. Montrer que pour tout plateau contenant au moins un carré 2×2, et pour tout $(x, y) \in \mathbb{F}_4 \times \mathbb{F}_4$, il existe une configuration S du plateau tel que $A(S) = x$ et $B(S) = y$.

e. On positionne la case centrale en $(0, 0)$.

(i) Calculer le couple $\big(A(S), B(S)\big)$ lorsque S est la configuration initiale du plateau de jeu français. Peut-on gagner ? Sur quelles cases doit-on enlever le pion initial pour espérer pouvoir gagner ?

(ii) Calculer le couple $\big(A(S), B(S)\big)$ lorsque S est la configuration initiale du plateau de jeu anglais.

Indication. Afin de faciliter les calculs, on pourra commencer par calculer $A(S)$ et $B(S)$ lorsque $S = \{(i, j) \mid i, j \in [\![-3, 3]\!]\}$, et utiliser la question **b**. On pourra aussi remarquant que $\alpha^3 = 1_{\mathbb{F}_4}$. Enfin, on pourra aussi montrer que trois billes sur des cases verticales consécutives n'influent pas sur la valeur de $A(S)$ et $B(S)$.

(iii) Montrer que sur un jeu de solitaire anglais résolu (i.e. ne contenant plus qu'un seul pion), le dernier pion est nécessairement situé sur une des cases

$$(0,-3), \quad (0,0), \quad (0,3), \quad (-3,0), \quad (3,0).$$

E16. Soit A un anneau commutatif, et soit $B = A[T]/(T^2)$. On pose $\varepsilon = \overline{X}$. On définit une loi externe

$$\begin{aligned} A \times B &\longrightarrow B \\ (a, \overline{P}) &\longmapsto a \cdot \overline{P} = \overline{a}\overline{P}. \end{aligned}$$

a. Montrer que tout élément de B s'écrit de manière unique sous la forme

$$a\overline{1} + b \cdot \varepsilon, \quad a, b \in A.$$

b. Calculer $B^\times$ et $\mathrm{Nil}(B)$.

c. Pour tout $P \in A[X]$, et tout $a \in A$, montrer que

$$P(a \cdot \overline{1} + \varepsilon) = P(a) \cdot \overline{1} + P'(a) \cdot \varepsilon.$$

E17. On reprend les notations de l'exercice XV-18, et l'on suppose que G est cyclique d'ordre n.

a. Montrer que pour tout anneau commutatif A, on a

$$A[G] \simeq A[X]/(X^n - 1).$$

b. On suppose que $n = 2$, et que $2 \in A^\times$. Montrer que $A[G] \simeq A \times A$.

c. Montrer que $\mathbb{C}[G] \simeq \mathbb{C}^n$.

d. Montrer que

$$\mathbb{R}[G] \simeq \mathbb{R} \times \mathbb{C}^{\frac{n-1}{2}}$$

si n est impair, et que

$$\mathbb{R}[G] \simeq \mathbb{R}^2 \times \mathbb{C}^{\frac{n-2}{2}}$$

si n est pair.

e. Montrer que $\mathbb{Q}[G] \simeq \mathbb{Q} \times \mathbb{Q}[j]$ si $n = 3$ et $\mathbb{Q}[G] \simeq \mathbb{Q}^2 \times \mathbb{Q}[i]$ si $n = 4$.

E18.

a. Montrer que les idéaux $(X-1)$ et (X^2+X+1) sont comaximaux dans l'anneau $\mathbb{Q}[X]$.

b. Montrer que l'application

$$\begin{aligned} \varphi : \mathbb{Z}[X] &\longrightarrow \mathbb{Z}/3\mathbb{Z} \\ P &\longmapsto \overline{P(1)} \end{aligned}$$

induit un isomorphisme d'anneaux $\mathbb{Z}[X]/(X-1,3) \longrightarrow \mathbb{Z}/3\mathbb{Z}$.

c. Montrer que $(X-1,3) = (X-1, X^2+X+1)$ dans $\mathbb{Z}[X]$.

d. Les idéaux $(X-1)$ et (X^2+X+1) sont-ils comaximaux dans $\mathbb{Z}[X]$?

e. Montrer que le morphisme $\mathbb{Z}[X] \longrightarrow \mathbb{Z}[X]/(X-1) \times \mathbb{Z}[X]/(X^2+X+1)$ n'est pas surjectif.

E19. Soit A un anneau commutatif, et soit $a \in A$.

a. Montrer que l'on a un isomorphisme d'anneaux

$$A[X]/(X-a) \simeq A.$$

b. On suppose dans cette question que $a \in A^\times$. Établir un isomorphisme d'anneaux

$$A[X]/(X(X-a)) \simeq A \times A.$$

c. Montrer qu'il existe quatre morphismes d'anneaux

$$\mathbb{Z}[X]/(X(X-2)) \longrightarrow \mathbb{Z} \times \mathbb{Z}.$$

Pour chaque morphisme, on explicitera l'image de $\overline{P} \in \mathbb{Z}[X]/\big(X(X-2)\big)$. Montrer alors qu'aucun d'eux n'est un isomorphisme.

Indication. On pourra montrer que si φ est un tel morphisme, on a alors $\mathrm{Im}(\varphi) \subset \{(a,b) \in \mathbb{Z} \times \mathbb{Z} \mid a \equiv b\ [2]\}$.

d. Si B est un anneau commutatif, on note $E(B)$ l'ensemble de ses idempotents.

Montrer qu'un isomorphisme d'anneaux $A_1 \simeq A_2$ induit une bijection entre $E(A_1)$ et $E(A_2)$.

e. Montrer que pour tout $\overline{P} \in A[X]/\big(X(X-a)\big)$, il existe $\alpha, \beta \in A$ **uniques** tels que $\overline{P} = \overline{\alpha + \beta X}$.

Attention ! L'anneau n'est pas supposé intègre !

On suppose que A est **intègre** et que $a \in A$ n'est **pas** inversible. En vous aidant de la question précédente, calculer $E(A[X]/\big(X(X-a)\big)$.

L'anneau $A[X]/\big(X(X-a)\big)$ est-il isomorphe à $A \times A$?

f. Soit A un anneau **intègre**. Donner une condition nécessaire et suffisante sur $a \in A$ pour que les anneaux $A[X]/\big(X(X-a)\big)$ et $A \times A$ soient isomorphes.

E20.

a. Soit $f : A \xrightarrow{\sim} B$ un isomorphisme d'anneaux. Montrer que pour tout idéal $\mathfrak{a}$ de A, $f(\mathfrak{a})$ est un idéal de B et que l'on a un isomorphisme

$$A/\mathfrak{a} \simeq B/f(\mathfrak{a}).$$

b. Soit A un anneau, et soient $\mathfrak{a}$ et $\mathfrak{b}$ deux idéaux de A. On note

$$\pi_{\mathfrak{a}} : A \longrightarrow A/\mathfrak{a} \quad \text{et} \quad \pi_{\mathfrak{b}} : A \longrightarrow A/\mathfrak{b}$$

les projections canoniques correspondantes.

Justifier que $\pi_{\mathfrak{a}}(\mathfrak{b})$ est un idéal de $A/\mathfrak{a}$, que $\pi_{\mathfrak{b}}(\mathfrak{a})$ est un idéal de $A/\mathfrak{b}$, et que l'on a des isomorphismes

$$(A/\mathfrak{a})/\pi_{\mathfrak{a}}(\mathfrak{b}) \simeq A/(\mathfrak{a},\mathfrak{b}) \simeq (A/\mathfrak{b})/\pi_{\mathfrak{b}}(\mathfrak{a}).$$

Indication. On pourra remarquer que $\pi_{\mathfrak{a}}(\mathfrak{b}) = (\mathfrak{a},\mathfrak{b})/\mathfrak{a}$.

E21. Soit $d \in \mathbb{Z} \setminus \{0,1\}$ un entier sans facteurs carrés, et soit $A = \mathbb{Z}[\sqrt{d}]$. Le but de cet exercice est de démontrer que tout idéal premier non nul de A est maximal.

Soit $\mathfrak{p}$ un idéal premier non nul de A.

a. Montrer que $\mathfrak{p} \cap \mathbb{Z}$ est un idéal premier de $\mathbb{Z}$.

b. Soit $z = a+b\sqrt{d} \in \mathfrak{p} \setminus \{0\}$. En considérant $(a+b\sqrt{d})(a-b\sqrt{d})$, montrer que $\mathfrak{p} \cap \mathbb{Z}$ est non nul.

c. En déduire qu'il existe un nombre premier p tel que $\mathfrak{p} \cap \mathbb{Z} = p\mathbb{Z}$.

d. Montrer que l'application

$$\begin{aligned} \mathbb{F}_p \times A/\mathfrak{p} &\longrightarrow A/\mathfrak{p} \\ (\widetilde{m}, \overline{z}) &\longmapsto \widetilde{m} * \overline{z} = \overline{mz} \end{aligned}$$

est bien définie, et induit une structure de $\mathbb{F}_p$-espace vectoriel sur $A/\mathfrak{p}$.

e. Montrer que $\dim_{\mathbb{F}_p}(A/\mathfrak{p}) \leqslant 2$.

f. Soit $\overline{z} \in A/\mathfrak{p}$ non nul. Montrer que l'application de multiplication par $\overline{z}$ dans $A/\mathfrak{p}$ est une application $\mathbb{F}_p$-linéaire injective. En déduire que $\overline{z}$ est inversible dans $A/\mathfrak{p}$.

g. En déduire que $\mathfrak{p}$ est maximal.

E22. Soit A un anneau commutatif non trivial.

a. On dit qu'une famille $\mathscr{F}'$ d'idéaux de A est *stable par réunion croissante* si pour toute famille d'éléments $(\mathfrak{a}_i)_{i\in I}$ de $\mathscr{F}'$ totalement ordonnée[4] par $(I,\leqslant)$, on a $\bigcup_{i\in I} \mathfrak{a}_i \in \mathscr{F}'$.

Eu utilisant le lemme de Zorn, montrer qu'une famille $\mathscr{F}'$ non vide d'idéaux de A stable par réunion croissante admet un élément maximal pour l'inclusion.

4. C'est-à-dire, pour tous $i, j \in I$, $i \leqslant j \implies \mathfrak{a}_i \subset \mathfrak{a}_j$.

b. Soit $\mathfrak{a}$ un idéal de A et soit $a \in A$. Montrer que l'ensemble

$$(\mathfrak{a} : a) = \{x \in A \mid ax \in \mathfrak{a}\}$$

est un idéal de A contenant $\mathfrak{a}$, et que l'on a

$$(\mathfrak{a}, a)(\mathfrak{a} : a) \subset \mathfrak{a}.$$

Ici, $(\mathfrak{a}, a)$ est l'idéal engendré par $\mathfrak{a}$ et a, i.e. $(\mathfrak{a}, a) = \mathfrak{a} + Aa$.

c. Soit $\mathscr{F}$ une famille d'idéaux de A vérifiant les conditions suivantes :

(O1) on a $A \in \mathscr{F}$;

(O2) pour tout idéal $\mathfrak{a}$, et tout $a \in A$, on a

$$(\mathfrak{a}, a) \in \mathscr{F} \text{ et } (\mathfrak{a} : a) \in \mathscr{F} \implies \mathfrak{a} \in \mathscr{F}.$$

On suppose que $\mathscr{F}' = \{\mathfrak{a} \text{ idéal de } A \mid \mathfrak{a} \notin \mathscr{F}\}$ est non vide, et possède un élément maximal $\mathfrak{p}$. Montrer que $\mathfrak{p}$ est un idéal premier.

Indication. Remarquer que $\mathfrak{p} \neq A$. Supposer qu'il existe a, $b \in A \setminus \mathfrak{p}$ tels que $ab \in \mathfrak{p}$, et montrer que les idéaux $(\mathfrak{p}, a)$ et $(\mathfrak{p} : a)$ contiennent strictement $\mathfrak{p}$.

d. Soit $\mathscr{F}$ une famille **non vide** d'idéaux de A vérifiant les conditions suivantes :

(F1) pour tous $\mathfrak{a}$, $\mathfrak{b} \in \mathscr{F}$, on a $\mathfrak{ab} \in \mathscr{F}$;

(F2) pour tous idéaux $\mathfrak{a}$, $\mathfrak{b}$ de A, on a

$$\mathfrak{a} \supset \mathfrak{b} \text{ et } \mathfrak{b} \in \mathscr{F} \implies \mathfrak{a} \in \mathscr{F}.$$

Montrer que $\mathscr{F}$ vérifie (O1) et (O2).

e. Soit S une partie de A non vide stable par multiplication.

(i) On suppose dans cette question que S ne contient pas 0, et l'on pose

$$\mathscr{F} = \{\mathfrak{a} \text{ idéal de } A \mid \mathfrak{a} \cap S \neq \emptyset\}.$$

Montrer que $\mathscr{F}$ vérifie (F1) et (F2), et que

$$\mathscr{F}' = \{\mathfrak{a} \text{ idéal de } A \mid \mathfrak{a} \notin \mathscr{F}\}$$

est non vide et stable par réunion croissante. En déduire qu'il existe un idéal premier de A disjoint de S.

(ii) Soit $\mathfrak{a}$ un idéal de A disjoint de S. En utilisant la question précédente, montrer que $\mathfrak{a}$ est contenu dans un idéal premier de A disjoint de S.

f. En considérant une famille $\mathscr{F}$ d'idéaux de A bien choisie, redémontrer que tout idéal maximal de A est premier.

g. Soit $\mathscr{F}$ une famille d'idéaux de A vérifiant (O1) et (O2), et telle que la famille $\mathscr{F}' = \{\mathfrak{a} \text{ idéal de } A \mid \mathfrak{a} \notin \mathscr{F}\}$ soit non vide et stable par réunion croissante.

Montrer que si $\mathscr{F}$ contient tous les idéaux premiers (resp. tous les idéaux premiers non nuls) de A, alors $\mathscr{F}$ contient tous les idéaux (resp. tous les idéaux non nuls) de A.

Indication. Considérer un idéal (resp. un idéal non nul) $\mathfrak{b} \notin \mathscr{F}$. Montrer que

$$\mathscr{F}'' = \{\mathfrak{a} \text{ idéal de } A \mid \mathfrak{a} \notin \mathscr{F},\ \mathfrak{b} \subset \mathfrak{a}\}$$

possède un élément maximal, et que c'est un élément maximal de $\mathscr{F}'$.

h. On dit que $a \in A$ est *régulier* si l'application

$$\begin{aligned} \ell_a \colon A &\longrightarrow A \\ x &\longmapsto ax \end{aligned}$$

est injective.

(i) Justifier que l'ensemble S des éléments réguliers de A est stable par multiplication, et ne contient pas 0.

(ii) En déduire qu'un idéal maximal parmi les idéaux ne contenant aucun élément régulier est un idéal premier.

(iii) Montrer que A est intègre si, et seulement si, tout idéal premier non nul de A contient un élément régulier.

E23. Soit A un anneau commutatif. Dans cet exercice, on propose une autre démonstration du fait que $\mathrm{Nil}(A)$ est l'intersection des idéaux premiers de A (cf. théorème 2.5). L'inclusion $\mathrm{Nil}(A) \subset \bigcap_{\mathfrak{p}} \mathfrak{p}$ se démontrant aisément, on se concentre sur l'inclusion plus difficile à établir.

a. Soit $\alpha \in A$. Sans utiliser la proposition XV-6.37, démontrer que le polynôme $\alpha X + 1 \in A[X]$ est inversible si, et seulement si, $\alpha \in \mathrm{Nil}(A)$.

Indication. Si $P \in A[X]$ vérifie $(\alpha X + 1)P = 1$, écrire

$$P = a_n X^n + \cdots + a_1 X + a_0,$$

et identifier les coefficients.

b. On suppose que $\alpha \in A \setminus \mathrm{Nil}(A)$.

(i) Justifier l'existence d'un idéal maximal $\mathfrak{m}$ de $A[X]$ contenant $\alpha X+1$.

(ii) Soit $\mathfrak{p} = \mathfrak{m} \cap A$. Montrer que $\mathfrak{p}$ est un idéal premier de A ne contenant pas α.

Indication. Pour montrer que $\alpha \in A \setminus \mathfrak{p}$, procéder par l'absurde.

c. Conclure.

Chapitre XVII

L'anneau $\mathbb{Z}/n\mathbb{Z}$ en long, en large et en travers

Dans ce chapitre, on étudie d'un peu plus près l'anneau $\mathbb{Z}/n\mathbb{Z}$. La première partie contient essentiellement des résultats qui ont déjà été vus dans les chapitres précédents ou les exercices.

1. Échauffement

Commençons par retraduire le théorème chinois dans le cas de l'anneau $\mathbb{Z}$. Pour tout $m \geqslant 1$ et tout $a \in \mathbb{Z}$, on note $[a]_m$ la classe de a modulo m. Le théorème XVI-3.13 et l'exemple XVI-3.11 nous donnent alors immédiatement le résultat suivant.

1.1. Théorème. *Soient $m_1, \ldots, m_r \geqslant 1$ des entiers premiers entre eux deux à deux. Alors, le morphisme*

$$\begin{aligned} f : \mathbb{Z} &\longrightarrow \mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_r\mathbb{Z} \\ a &\longmapsto \quad ([a]_{m_1}, \ldots, [a]_{m_r}) \end{aligned}$$

est surjectif, de noyau $(m_1 \times \cdots \times m_r)$. En particulier, on a un isomorphisme d'anneaux

$$\mathbb{Z}/(m_1 \times \cdots \times m_r)\mathbb{Z} \simeq \mathbb{Z}/m_1\mathbb{Z} \times \cdots \times \mathbb{Z}/m_r\mathbb{Z}.$$

1.2. Corollaire. *Soit $n \geqslant 2$, et soit $n = p_1^{m_1} \times \cdots \times p_r^{m_r}$ sa décomposition en facteurs premiers. Alors, le morphisme*

$$\begin{aligned} f : \mathbb{Z} &\longrightarrow \mathbb{Z}/p_1^{m_1} \times \cdots \times \mathbb{Z}/p_r^{m_r}\mathbb{Z}\mathbb{Z} \\ a &\longmapsto \quad ([a]_{p_1^{m_1}}, \ldots, [a]_{p_r^{m_r}}) \end{aligned}$$

induit un isomorphisme d'anneaux

$$\mathbb{Z}/n\mathbb{Z} \simeq \mathbb{Z}/p_1^{m_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_r^{m_r}\mathbb{Z}.$$

On détermine maintenant les éléments inversibles, les éléments nilpotents et les diviseurs de zéro de l'anneau $\mathbb{Z}/n\mathbb{Z}$.

1.3. Proposition. *Soit $n \geqslant 2$ et soit $a \in \mathbb{Z}$. Alors :*

(1) *$\overline{a}$ est inversible si, et seulement si, a et n sont premiers entre eux ;*

(2) *$\overline{a}$ est un diviseur de zéro si, et seulement si, $n \nmid a$ et a et n ne sont pas premiers entre eux ;*

(3) *$\overline{a}$ est nilpotent si, et seulement si, tout diviseur premier p de n est aussi un diviseur premier de a.*

Démonstration. Soit $a \in \mathbb{Z}$. Alors, on a la série d'équivalences suivantes :

$$\begin{aligned} \overline{a} \in (\mathbb{Z}/n\mathbb{Z})^\times &\iff \text{il existe } u \in \mathbb{Z} \text{ tel que } \overline{u}\,\overline{a} = \overline{1} \\ &\iff \text{il existe } u \in \mathbb{Z}, \text{ il existe } v \in \mathbb{Z} \text{ tels que } 1 - ua = vn \\ &\iff \text{il existe } u, v \in \mathbb{Z} \text{ tels que } ua + vn = 1 \\ &\iff a \text{ et } b \text{ sont premiers entre eux.} \end{aligned}$$

Supposons maintenant que $\overline{a}$ soit un diviseur de 0. Alors, $\overline{a} \neq \overline{0}$, i.e. $n \nmid a$. De plus, il existe $\overline{b} \neq \overline{0}$ tel que $\overline{a}\overline{b} = \overline{0}$. On a donc $n \mid ab$. Si a et n étaient premiers entre eux, alors on aurait $n \mid b$ par le lemme de Gauss, d'où la contradiction $\overline{b} = \overline{0}$. Ainsi, a et n ne sont pas premiers entre eux.

Réciproquement, supposons que $n \nmid a$ et que a et n ne soient pas premiers entre eux. En particulier, $\overline{a} \neq \overline{0}$, et il existe un nombre premier p divisant à la fois a et n. On peut donc écrire $a = pa'$ et $n = pn'$, avec $a', n' \in \mathbb{Z}$. Remarquons que, puisque $p \geqslant 2$, on a $n' \in [\![1, n-1]\!]$. En particulier, on a $\overline{n'} \neq \overline{0}$. Mais alors, on obtient

$$\overline{a}\overline{n'} = \overline{an'} = \overline{pa'n'} = \overline{a'n} = \overline{0}.$$

Ainsi, $\overline{a}$ est un diviseur de 0.

Supposons enfin que $\overline{a}$ soit nilpotent. Il existe donc $m \geqslant 1$ tel que $\overline{a}^m = \overline{0}$. Cela revient à dire que $n \mid a^m$. Soit p un diviseur premier de n. On a alors $p \mid a^m$, et par le lemme d'Euclide, $p \mid a$. Inversement, supposons que tout diviseur premier de n divise aussi a, et soit $n = p_1^{m_1} \cdots p_r^{m_r}$ la décomposition en facteurs premiers de n. Soit $m = \max(m_1, \ldots, m_r)$. Par hypothèse, a est divisible par chaque p_i, donc par $p_1 \times \cdots \times p_r$. Ainsi, a^m est divisible par $(p_1 \cdots p_r)^m$. Mais ce dernier entier est un multiple de n par choix de m, et ainsi $\overline{a}^m = \overline{a^m} = \overline{0}$. Ainsi, $\overline{a}$ est nilpotent. Cela achève la démonstration. □

1.4. Remarque. Cette proposition permet de retrouver immédiatement le résultat de la proposition XV-3.9.

La détermination de la structure du groupe $\mathrm{Nil}(\mathbb{Z}/n\mathbb{Z})$ est proposée en exercice au lecteur.

Passons maintenant aux éléments idempotents.

1.5. Proposition. *Soit $n \geqslant 2$, et soit r le nombre de diviseurs premiers distincts de n. Alors, l'anneau $\mathbb{Z}/n\mathbb{Z}$ possède exactement 2^r idempotents. En particulier, pour tout nombre premier p, et tout $m \geqslant 1$, $\mathbb{Z}/p^m\mathbb{Z}$ n'admet pas d'idempotents non triviaux.*

Démonstration. Soit $n = p_1^{m_1} \times \cdots \times p_r^{m_r}$ la décomposition en facteurs premiers de n, et soit

$$\begin{array}{rcl} \varphi : \mathbb{Z}/n\mathbb{Z} & \xrightarrow{\sim} & \mathbb{Z}/p_1^{m_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_r^{m_r}\mathbb{Z} \\ [a]_n & \longmapsto & ([a]_{p_1^{m_1}}, \ldots, [a]_{p_r^{m_r}}) \end{array}$$

l'isomorphisme donné par le théorème chinois. Alors, on a

$$[a]_n^2 = [a]_n \iff \varphi([a]_n^2) = \varphi([a]_n) \iff \varphi([a]_n)^2 = \varphi([a]_n).$$

Ainsi, φ induit une bijection de l'ensemble des idempotents de $\mathbb{Z}/n\mathbb{Z}$ sur l'ensemble des idempotents de $\mathbb{Z}/p_1^{m_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_r^{m_r}\mathbb{Z}$.
De plus, l'élément $([a_1]_{p_1^{m_1}}, \ldots, [a_r]_{p_r^{m_r}}) \in \mathbb{Z}/p_1^{m_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_r^{m_r}\mathbb{Z}$ est un idempotent si, et seulement si, $[a_i]_{p_i^{m_i}}$ est un idempotent de $\mathbb{Z}/p_i^{m_i}\mathbb{Z}$, pour tout $i \in [\![1, r]\!]$.

Il reste donc à déterminer les idempotents de $\mathbb{Z}/p^m\mathbb{Z}$, lorsque p est un nombre premier et $m \geqslant 1$. Soit $[a]_{p^m} \in \mathbb{Z}/p^m\mathbb{Z}$. On peut toujours supposer que $a \in [\![0, p^m - 1]\!]$. Si $a = 0$, alors $\overline{a} = \overline{0}$ est bien un idempotent. Supposons maintenant que $a \geqslant 1$. On peut donc écrire

$$a = up^k, \quad \text{avec } p \nmid u \text{ et } k \in [\![0, m-1]\!].$$

Supposons que $[a]_{p^m}$ soit un idempotent. On a donc $[a(1-a)]_{p^m} = [0]_{p^m}$, i.e. $p^m \mid a(1-a)$, ou encore $p^{m-k} \mid u(1-up^k)$. Puisque $p \nmid u$, on en déduit que $p^{m-k} \mid 1 - up^k$. Si $r = \min(k, m-k)$, on a donc $p^r \mid 1$, ce qui est une contradiction, à moins que $r = 0$. On a donc $k = 0$ ou $k = m$. Si $k = m$, alors $\overline{a} = \overline{up^m} = \overline{0}$. Si $k = 0$, on a $a = u$, et ainsi $p \nmid a$. Comme $p^m \mid a(1-a)$, on en déduit que $p^m \mid 1 - a$, soit $\overline{a} = \overline{1}$.

Finalement, les seuls idempotents de $\mathbb{Z}/p^m\mathbb{Z}$ sont $\overline{0}$ et $\overline{1}$. Les considérations précédentes montrent alors que les idempotents de $\mathbb{Z}/n\mathbb{Z}$ sont les éléments

$$\varphi^{-1}\left(([a_1]_{p_1^{m_1}}, \ldots, [a_r]_{p_r^{m_r}})\right), \quad a_i \in \{0, 1\}.$$

Il y en a donc 2^r. Cela achève la démonstration. □

1.6. Remarque. D'après la fin de la démonstration précédente, dont on garde les notations, les idempotents de $\mathbb{Z}/n\mathbb{Z}$ sont les éléments

$$\varphi^{-1}\big(([a_1]_{p_1^{m_1}}, \ldots, [a_r]_{p_r^{m_r}})\big), \quad a_i \in \{0,1\}.$$

Les calculs faits à la fin du chapitre précédent montrent que l'isomorphisme inverse est donné par

$$\begin{aligned} \varphi^{-1} : \mathbb{Z}/p_1^{m_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_r^{m_r}\mathbb{Z} &\longrightarrow \mathbb{Z}/n\mathbb{Z} \\ ([a_1]_{p_1^{m_1}}, \ldots, [a_n]_{p_r^{m_r}}) &\longmapsto \sum_{i=1}^{r} [e_i a_i]_n, \end{aligned}$$

où $e_i = \varphi^{-1}([0]_{p_1^{m_1}}, \ldots, [0]_{p_{i-1}^{m_{i-1}}}, [1]_{p_i^{m_i}}, [0]_{p_{i+1}^{m_{i+1}}}, \ldots, [0]_{p_r^{m_r}})$.

Ainsi, les idempotents de $\mathbb{Z}/n\mathbb{Z}$ sont les éléments

$$[a_1 e_1 + \cdots + a_r e_r]_n, \quad a_i \in \{0,1\}.$$

2. Structure de $(\mathbb{Z}/n\mathbb{Z})^\times$

Nous allons dans ce paragraphe élucider complètement la structure du groupe $(\mathbb{Z}/n\mathbb{Z})^\times$. Commençons par calculer l'ordre de ce groupe. Pour tout entier $n \geqslant 1$, on rappelle que $\varphi(n)$ désigne le nombre d'entiers de $[\![1, n]\!]$ qui sont premiers avec n.

On a alors un premier résultat.

2.1. Lemme. *Pour tout $n \geqslant 2$, on a $|(\mathbb{Z}/n\mathbb{Z})^\times| = \varphi(n)$.*

Démonstration. Notons tout d'abord que chaque élément de $\mathbb{Z}/n\mathbb{Z}$ admet un unique représentant à valeur dans $[\![1, n]\!]$. Soit $a \in [\![1, n]\!]$. D'après la proposition 1.3 (1), on sait que $\overline{a} \in (\mathbb{Z}/n\mathbb{Z})^\times$ si, et seulement si, a est premier avec n. Le résultat découle alors de la définition de $\varphi(n)$. □

Puisqu'un isomorphisme d'anneaux $A \simeq B$ induit un isomorphisme de groupes $A^\times \simeq B^\times$, le théorème 1.1 et le corollaire 1.2 nous donnent en particulier le lemme suivant.

2.2. Lemme

(1) *Soient $n, m \geqslant 2$ deux entiers premiers entre eux. Alors, on a un isomorphisme de groupes*

$$(\mathbb{Z}/nm\mathbb{Z})^\times \simeq (\mathbb{Z}/n\mathbb{Z})^\times \times (\mathbb{Z}/m\mathbb{Z})^\times.$$

(2) *Soit $n \geqslant 2$ un entier, et soit $n = p_1^{m_1} \times \cdots \times p_r^{m_r}$ sa décomposition en facteurs premiers. Alors, on a un isomorphisme de groupes*

$$(\mathbb{Z}/n\mathbb{Z})^\times \simeq (\mathbb{Z}/p_1^{m_1}\mathbb{Z})^\times \times \cdots \times (\mathbb{Z}/p_r^{m_r}\mathbb{Z})^\times.$$

Cela permet de retrouver directement le résultat du corollaire VI-6.7, sans utiliser la formule d'inversion de Möbius.

2.3. Corollaire

(1) *Soient n, $m \geqslant 2$ deux entiers premiers entre eux. Alors, on a*

$$\varphi(nm) = \varphi(n)\varphi(m).$$

(2) *Pour tout nombre premier p, et tout entier $m \geqslant 1$, on a*

$$\varphi(p^m) = p^m - p^{m-1}.$$

(3) *Soit $n \geqslant 2$ un entier, et soit $n = p_1^{m_1} \times \cdots \times p_r^{m_r}$ sa décomposition en facteurs premiers. Alors, on a*

$$\varphi(n) = (p_1^{m_1} - p_1^{m_1-1}) \times \cdots \times (p_r^{m_r} - p_r^{m_r-1}).$$

Démonstration. Le point (1) découle des deux lemmes précédents. Pour établir (2), il suffit de constater que les entiers de l'ensemble $[\![1, p^m]\!]$ qui ne sont pas premiers avec p^m sont exactement les multiples de p. Il y en a donc p^{m-1}, d'où le résultat. Le point (3) est une application immédiate des deux premiers points. Cela achève la démonstration. □

Le lemme 2.2 (2) ci-dessus, nous montre que pour comprendre la structure du groupe $(\mathbb{Z}/n\mathbb{Z})^\times$, il suffit de comprendre celle du groupe $(\mathbb{Z}/p^m\mathbb{Z})^\times$, où p est premier, et $m \geqslant 1$. C'est l'objet du théorème suivant.

2.4. Théorème. *Soit p un nombre premier et soit $m \geqslant 1$.*

(1) *Si p est impair, $(\mathbb{Z}/p^m\mathbb{Z})^\times$ est cyclique d'ordre $p^{m-1}(p-1)$. De plus, si $a \in \mathbb{Z}$ est tel que $[a]_p$ engendre $(\mathbb{Z}/p\mathbb{Z})^\times$, alors $[a^{p^{m-1}}(p+1)]_{p^m}$ engendre $(\mathbb{Z}/p^m\mathbb{Z})^\times$.*

(2) *Le groupe $(\mathbb{Z}/2\mathbb{Z})^\times$ est trivial, et le groupe $(\mathbb{Z}/4\mathbb{Z})^\times$ est cyclique d'ordre 2, engendré par $[-1]_4$.*

(3) *Pour tout $m \geqslant 3$, $[-1]_{2^m}$ est d'ordre 2, $[5]_{2^m}$ est d'ordre 2^{m-2}, et l'on a la décomposition en produit direct*

$$(\mathbb{Z}/2^m\mathbb{Z})^\times = \langle [-1]_{2^m} \rangle \odot \langle [5]_{2^m} \rangle.$$

En particulier, on a un isomorphisme de groupes

$$(\mathbb{Z}/2^m\mathbb{Z})^\times \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{m-2}\mathbb{Z},$$

et $(\mathbb{Z}/2^m\mathbb{Z})^\times$ n'est pas cyclique.

Démonstration. Supposons que $p > 2$. D'après le corollaire XII-2.9, le groupe $(\mathbb{Z}/p\mathbb{Z})^\times$ est cyclique. Soit $[a]_p$ un générateur de $(\mathbb{Z}/p\mathbb{Z})^\times$. Nous allons tout d'abord démontrer que $[a^{p^{m-1}}]_{p^m}$ est d'ordre $p-1$.

Soit $d = o([a^{p^{m-1}}]_{p^m})$. Comme $(\mathbb{Z}/p^m\mathbb{Z})^\times$ est d'ordre $\varphi(p^m) = p^{m-1}(p-1)$, on a

$$[a^{p^{m-1}}]_{p^m}^{p-1} = [a]_{p^m}^{p^{m-1}(p-1)} = [1]_{p^m}$$

par le théorème de Lagrange. Ainsi, $d \mid p-1$.

De plus, l'égalité $[a^{p^{m-1}}]_{p^m}^d = [1]_{p^m}$ implique[1] que $[a^{p^{m-1}}]_p^d = [1]_p$. D'après le petit théorème de Fermat, on a $a^p \equiv a\ [p]$, soit $[a^p]_p = [a]_p$. Il en découle aisément l'égalité $[a]_p^d = [1]_p$. Mais alors, $o([a]_p) \mid d$, i.e. $p-1 \mid d$. On en déduit donc que $d = p-1$, ce qui était le résultat voulu.

Montrons maintenant que $o([p+1]_{p^m}) = p^{m-1}$. Pour cela, nous allons montrer la propriété suivante pour tout $k \geqslant 0$:

(H_k) Il existe $u_k \in \mathbb{Z}$, $p \nmid u_k$ tel que $(p+1)^{p^k} = 1 + u_k p^{k+1}$.

La propriété (H_0) est clairement vraie, puisqu'il suffit de poser $u_0 = 1$. Supposons que (H_k) soit vraie pour un certain entier $k \geqslant 0$, et montrons que (H_{k+1}) est vraie. Par hypothèse, il existe $u_k \in \mathbb{Z}, p \nmid u_k$ tel que

$$(p+1)^{p^k} = 1 + u_k p^{k+1}.$$

Mais alors, en utilisant la formule du binôme, on obtient

$$(p+1)^{p^{k+1}} = (1+u_k p^{k+1})^p = 1 + u_k p^{k+2} + \sum_{i=2}^{p-1} \binom{p}{i} u_k^i p^{i(k+1)} + u_k^p p^{p(k+1)}.$$

Or, pour tout $i \in [\![2, p-1]\!]$, $\binom{p}{i}$ est divisible par p. De plus, on a

$$i(k+1) + 1 \geqslant 2k+3 \geqslant k+3.$$

Ainsi, $\displaystyle\sum_{i=2}^{p-1} \binom{p}{i} u_k^i p^{i(k+1)}$ est divisible par p^{k+3}. Enfin, comme $p > 2$, on a aussi $p(k+1) \geqslant 3k+3 \geqslant k+3$.

Finalement, $(p+1)^{p^{k+1}} = 1 + u_k p^{k+2} + v_k p^{k+3}$ pour un certain entier $v_k \in \mathbb{Z}$. Il suffit donc de poser $u_{k+1} = u_k + p v_k$, en remarquant que, puisque $p \nmid u_k$, on a aussi $p \nmid u_{k+1}$.

Par principe de récurrence, (H_k) est vraie pour tout $k \geqslant 0$.

Soit maintenant $k \in [\![0, m-2]\!]$. Alors, on a $[p+1]_{p^m}^{p^k} \neq [1]_{p^m}$. En effet, dans le cas contraire, on aurait $p^m \mid (p+1)^{p^k} - 1$, soit $p^m \mid u_k p^{k+1}$, ou encore $p^{m-k-1} \mid u_k$. Mais alors, on aurait $p \mid u_k$, d'où une contradiction. En revanche, on a

$$[p+1]_{p^m}^{p^{m-1}} = [(p+1)^{p^{m-1}}]_{p^m} = [1 + u_{m-1} p^m]_{p^m} = [1]_{p^m}.$$

Tout cela démontre que $[p+1]_{p^m}$ est d'ordre p^{m-1}.

1. Il suffit de revenir à la définition de congruence pour le voir.

Comme les éléments $[a^{p^{m-1}}]_{p^m}$ et $[p+1]_{p^m}$ commutent et ont des ordres premiers entre eux, la proposition VI-4.17 (2) montre que l'on a

$$o([a^{p^{m-1}}(p+1)]_{p^m}) = o([a^{p^{m-1}}]_{p^m}[p+1]_{p^m}) = (p-1)p^{m-1} = \varphi(p^m).$$

Ainsi, $(\mathbb{Z}/p^m\mathbb{Z})^\times$ est cyclique, engendré par $[a^{p^{m-1}}(p+1)]_{p^m}$. Cela démontre le point (1).

Le point (2) étant un calcul immédiat, montrons (3). L'élément $[-1]_{2^m}$ est clairement d'ordre au plus 2, et comme $2^m \geqslant 4$, $[-1]_{2^m} \neq [1]_{2^m}$ et $[-1]_{2^m}$ est d'ordre 2.

Montrons que $[5]_{2^m}$ est d'ordre 2^{m-2}. On va montrer par récurrence la propriété suivante pour tout $k \geqslant 0$:

(H'_k) Il existe un entier impair $u_k \in \mathbb{Z}$, tel que $5^{2^k} = 1 + u_k 2^{k+2}$.

La propriété (H'_0) est vraie, comme on le voit en posant $u_0 = 1$. Supposons que (H'_k) soit vraie pour un entier $k \geqslant 0$, et montrons que (H'_{k+1}) est vraie. Par hypothèse, il existe $u_k \in \mathbb{Z}$ impair tel que $5^{2^k} = 1 + u_k 2^{k+2}$. On a alors

$$5^{2^{k+1}} = (1 + u_k 2^{k+2})^2 = 1 + u_k 2^{k+3} + u_k^2 2^{2(k+2)} = 1 + (u_k + 2^{k+1}u_k^2)2^{k+3}.$$

Il suffit donc de poser $u_{k+1} = u_k + 2^{k+1}u_k^2$ et de vérifier que u_{k+1} est impair pour achever la récurrence.

On en déduit alors[2] comme précédemment que $o(5) = 2^{m-2}$. Nous allons maintenant établir la décomposition en produit direct

$$(\mathbb{Z}/2^m\mathbb{Z})^\times = \langle [-1]_{2^m} \rangle \odot \langle [5]_{2^m} \rangle.$$

Commençons par montrer que les deux sous-groupes considérés s'intersectent trivialement.

Soit $[a]_{2^m} \in \langle [-1]_{2^m} \rangle \cap \langle [5]_{2^m} \rangle$. Il existe donc $k \in [\![0, m-3]\!]$ et $\ell \in \{0,1\}$ tels que

$$[a]_{2^m} = [5]_{2^m}^k = [-1]_{2^m}^\ell.$$

En particulier,

$$5^k \equiv (-1)^\ell \ [2^m].$$

Comme $m \geqslant 3$, cela implique que l'on a $1 \equiv (-1)^\ell$ [4]. Cela n'est possible que si $\ell = 0$, et par conséquent $[a]_{2^m} = [1]_{2^m}$.

D'après ce qui précède, les sous-groupes $\langle [-1]_{2^m} \rangle$ et $\langle [5]_{2^m} \rangle$ sont respectivement d'ordre 2 et 2^{m-2}. Comme on travaille dans un groupe abélien, on obtient alors la décomposition voulue en utilisant le corollaire X-2.9. L'isomorphisme annoncé découle alors immédiatement de cette décomposition.

2. Cela est aussi valable en fait pour $m = 2$.

Enfin, un calcul immédiat dans le groupe $\mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2^{m-2}\mathbb{Z}$ montre que l'ordre d'un élément de $(\mathbb{Z}/2^m\mathbb{Z})^\times$ est au plus[3] 2^{m-2}. En particulier, le groupe $(\mathbb{Z}/2^m\mathbb{Z})^\times$ n'a pas d'élément d'ordre 2^{m-1}, et n'est donc pas cyclique. Cela achève la démonstration. □

2.5. Remarques

(1) La décomposition en produit direct interne et l'isomorphisme énoncés dans le point (3) sont valables même si $m = 2$. Dans ce cas, le second facteur est le groupe trivial. En revanche, la non-cyclicité de $(\mathbb{Z}/2^m\mathbb{Z})^\times$ n'est valable que pour $m \geqslant 3$.

(2) Le point (1) montre que l'on sait expliciter un générateur de $(\mathbb{Z}/p^m\mathbb{Z})^\times$ dès qu'on sait le faire pour $m = 1$. Ce problème est extrêmement difficile en général, mais permet de trouver à la main des résultats intéressants pour de petits nombres premiers.
Par exemple, puisque $[-1]_3$ engendre $(\mathbb{Z}/3\mathbb{Z})^\times$, on obtient que $[-4]_{3^m}$ engendre $(\mathbb{Z}/3^m\mathbb{Z})^\times$ pour tout $m \geqslant 1$.

3. Carrés de $\mathbb{Z}/n\mathbb{Z}$

Nous allons maintenant nous intéresser aux carrés de l'anneau $\mathbb{Z}/n\mathbb{Z}$.

Notation. Si $n \geqslant 2$ est un entier, on note[4]

$$(\mathbb{Z}/n\mathbb{Z})^2 = \{[a]_n^2 \mid [a]_n \in \mathbb{Z}/n\mathbb{Z}\},$$

ainsi que

$$(\mathbb{Z}/n\mathbb{Z})^{\times 2} = \{[a]_n^2 \mid [a]_n \in (\mathbb{Z}/n\mathbb{Z})^\times\}.$$

Le théorème chinois donne immédiatement le lemme suivant.

3.1. Lemme. *Soit $n \geqslant 2$ un entier, soit $n = p_1^{m_1} \times \cdots \times p_r^{m_r}$ sa décomposition en facteurs premiers et soit $[a]_n \in \mathbb{Z}/n\mathbb{Z}$ (resp. $(\mathbb{Z}/n\mathbb{Z})^\times$).*

Alors, on a $[a]_n \in (\mathbb{Z}/n\mathbb{Z})^2$ (resp. $[a]_n \in (\mathbb{Z}/n\mathbb{Z})^{\times 2}$) si, et seulement si, on a $[a]_{p_i^{m_i}} \in (\mathbb{Z}/p_i^{m_i}\mathbb{Z})^2$ (resp. $[a]_{p_i^{m_i}} \in (\mathbb{Z}/p_i^{m_i}\mathbb{Z})^{\times 2}$), pour tout $i \in [\![1, r]\!]$.

En particulier, on a

$$|(\mathbb{Z}/n\mathbb{Z})^2| = |(\mathbb{Z}/p_1^{m_1}\mathbb{Z})^2| \times \cdots \times |(\mathbb{Z}/p_r^{m_r}\mathbb{Z})^2|$$

et

$$|(\mathbb{Z}/n\mathbb{Z})^{\times 2}| = |(\mathbb{Z}/p_1^{m_1}\mathbb{Z})^{\times 2}| \times \cdots \times |(\mathbb{Z}/p_r^{m_r}\mathbb{Z})^{\times 2}|.$$

Il suffit donc de s'intéresser aux carrés de $\mathbb{Z}/p^m\mathbb{Z}$ lorsque p est un nombre

3. Cela utilise le fait que $m \geqslant 3$.

4. L'auteur se fait un devoir de faire en sorte qu'il n'y ait pas de confusion possible avec le produit direct de deux copies de $\mathbb{Z}/n\mathbb{Z}$ ou $(\mathbb{Z}/n\mathbb{Z})^\times$ dans ce chapitre.

premier et $m \geqslant 1$. Le but des résultats qui suivent est de démontrer qu'il suffit de se réduire à la détermination des carrés de $\mathbb{Z}/p\mathbb{Z}$ lorsque p est impair.

En vue d'une application ultérieure, nous allons tout d'abord déterminer les entiers $a \in \mathbb{Z}$ qui sont des carrés modulo p^m pour **tout** $m \geqslant 1$. On commence par le cas des entiers premiers avec p.

3.2. Lemme. *Soit $a \in \mathbb{Z}$, et soit p un nombre premier tel que $p \nmid a$.*

(1) *Supposons que $p = 2$ et que $a \equiv 1$ [8]. Alors, pour tout $m \geqslant 1$, l'équation*
$$x^2 \equiv a \ [p^m]$$
admet une solution.

(2) *Supposons que $p > 2$, et que a soit un carré modulo p. Alors, pour tout entier $m \geqslant 1$, l'équation*
$$x^2 \equiv a \ [p^m]$$
admet une solution.

Démonstration

On procède par récurrence dans les deux cas. Supposons d'abord que $p = 2$, et écrivons $a = 1 + 8s$. Alors, on a $1^2 \equiv a \ [2^m]$, $m = 1, 2, 3$.

Supposons maintenant que l'on ait l'existence d'une solution x_m de l'équation $x^2 \equiv a \ [2^m]$ pour un entier $m \geqslant 3$. On cherche une solution x_{m+1} de l'équation $x^2 \equiv a \ [2^{m+1}]$ de la forme $x_{m+1} = x_m + 2^{m-1}\lambda$, où $\lambda \in \mathbb{Z}$ reste encore à déterminer.

On a donc
$$x_{m+1}^2 = x_m^2 + 2^m \lambda x_m + 2^{2m-2}\lambda^2.$$

Or, $m \geqslant 3$, donc $2m - 2 \geqslant m + 1$, et l'on a ainsi
$$x_{m+1}^2 \equiv x_m^2 + 2^m \lambda x_m \ [2^{m+1}].$$

Comme a est impair, x_m aussi, et l'on a alors
$$x_{m+1}^2 \equiv x_m^2 + 2^m \lambda \ [2^{m+1}].$$

Par hypothèse, on a $x_m^2 \equiv a \ [2^m]$. On a donc $x_m^2 \equiv a$ ou $a + 2^m \ [2^{m+1}]$. Il suffit alors de prendre $\lambda = 0$ dans le premier cas, et $\lambda = 1$ dans le second cas pour obtenir la solution désirée. Cela achève la récurrence dans ce cas.

Supposons maintenant que $p > 2$. Par hypothèse, l'équation $x^2 \equiv a \ [p]$ a une solution x_1.

On suppose qu'il existe une solution x_m à l'équation $x^2 \equiv a\ [p^m]$ pour un entier $m \geqslant 1$. On cherche une solution x_{m+1} de l'équation $x^2 \equiv a\ [p^{m+1}]$ de la forme $x_{m+1} = x_m + p^m\lambda$, où $\lambda \in \mathbb{Z}$ reste à déterminer. On a donc

$$x_{m+1}^2 = x_m^2 + 2p^m\lambda x_m + p^{2m}\lambda^2.$$

Comme $m \geqslant 1$, on a $2m \geqslant m+1$, et l'on a

$$x_{m+1}^2 \equiv x_m^2 + 2p^m\lambda x_m\ [p^{m+1}].$$

Par hypothèse, on a $x_m^2 \equiv a\ [p^m]$. On peut donc écrire $x_m^2 = a + p^m k,\ k \in \mathbb{Z}$. On a ainsi

$$x_{m+1}^2 = a + p^m(2\lambda x_m + k)\ [p^{m+1}].$$

Comme $p \nmid a$, on a aussi $p \nmid x_m$. Comme p est impair, $2x_m$ est alors inversible modulo p. Soit y un inverse de $2x_m$ modulo p, et posons $\lambda = -ky$. Alors, on a $2\lambda x_m + k \equiv 0\ [p]$, et l'on obtient par conséquent une solution x_{m+1} de l'équation $x^2 \equiv a\ [p^{m+1}]$.

Cela achève la démonstration. □

On peut maintenant passer au cas général.

3.3. Lemme. *Soit p un nombre premier et soit $a \in \mathbb{Z}$ un entier non nul. Écrivons $a = p^\ell a_0$, avec $\ell \geqslant 0$ et $p \nmid a_0$.*

(1) *L'équation $x^2 \equiv a\ [2^m]$ a une solution pour tout $m \geqslant 1$ si, et seulement si, ℓ est pair et $a_0 \equiv 1\ [8]$.*

(2) *L'équation $x^2 \equiv a\ [p^m]$ a une solution pour tout $m \geqslant 1$ si, et seulement si, ℓ est pair et a_0 est un carré modulo p.*

Démonstration. Soit p un nombre premier [5]. Si $x \in \mathbb{Z}\setminus\{0\}$, on note $v_p(x)$ sa valuation p-adique. Supposons que l'équation $x^2 \equiv a\ [p^m]$ ait une solution pour tout $m \geqslant 1$. Alors, l'équation

$$x^2 \equiv a\ [p^{\ell+3}]$$

possède une solution. Quitte à rajouter un multiple de $p^{\ell+3}$, on peut supposer x non nul. Si $x \in \mathbb{Z}$ est une telle solution, on a donc

$$x^2 = p^\ell a_0 + kp^{\ell+3} = p^\ell(a_0 + kp^3)$$

pour un certain $k \in \mathbb{Z}$.

Comme $p \nmid a_0$, on a

$$v_p(x^2) = 2v_p(x) = v_p(p^\ell a_0 + kp^{\ell+3}) = \ell.$$

5. Pas nécessairement impair.

Ainsi, ℓ est pair. On peut alors écrire

$$x = p^{\frac{\ell}{2}} x_0, \quad p \nmid x_0,$$

et en simplifiant alors par p^ℓ, on obtient $x_0^2 = a_0 + kp^3$, d'où

$$x_0^2 \equiv a_0 \ [p^3].$$

Si $p = 2$, on en déduit que a_0 est un carré modulo 8, et donc que $a_0 \equiv 1$ [8] (car a_0 est impair). Si $p > 2$, on en déduit que a_0 est un carré modulo p^3, donc un carré modulo p. Les conditions du lemme sont donc nécessaires.

Montrons maintenant qu'elles sont suffisantes. Par le lemme 3.2, a_0 est un carré modulo p^m pour tout $m \geqslant 1$. Puisque ℓ est pair, p^ℓ est également un carré modulo p^m pour tout $m \geqslant 1$. On en déduit alors aisément que a est un carré modulo p^m pour tout $m \geqslant 1$. □

On revient maintenant à notre question initiale en déterminant tous les carrés de $\mathbb{Z}/p^m\mathbb{Z}$ à m fixé.

3.4. Proposition. *Soit p un nombre premier, soit $a \in \mathbb{Z}$ un entier non nul et soit $m \geqslant 1$. Écrivons $a = p^\ell a_0$, avec $\ell \geqslant 0$ et $p \nmid a_0$.*

(1) *Si $p = 2$, alors a est un carré modulo p^m si, et seulement si, une des conditions suivantes est vérifiée :*

   (i) $\ell \geqslant m$;

   (ii) *ℓ est pair et $m = \ell + 1$;*

   (iii) *ℓ est pair, $m = \ell + 2$ et $a_0 \equiv 1$ [4] ;*

   (iv) *ℓ est pair, $m \geqslant \ell + 3$ et $a_0 \equiv 1$ [8]. Dans ce dernier cas, a est un carré modulo p^k, pour tout $k \geqslant 1$.*

(2) *Si $p > 2$, alors a est un carré modulo p^m si, et seulement si, une des conditions suivantes est vérifiée :*

   (i) $\ell \geqslant m$;

   (ii) *$\ell < m$, ℓ est pair et a_0 est un carré modulo p. Dans ce dernier cas, a est un carré modulo p^k, pour tout $k \geqslant 1$.*

Démonstration. Soit p un nombre premier. Si $\ell \geqslant m$, a est nul modulo p^m et est un carré modulo p^m.

On suppose maintenant que $\ell < m$ jusqu'à la fin de la démonstration. Supposons tout d'abord que a soit un carré modulo p^m. Il existe donc un entier $x \in \mathbb{Z}$ tel que $x^2 \equiv a \ [p^m]$.

Quitte à rajouter un multiple de p^m, on peut supposer l'entier x non nul. Écrivons $x = p^s x_0$, $p \nmid x_0$. Il existe alors $y \in \mathbb{Z}$ tel que

$$x^2 = a + yp^m.$$

Puisque $\ell = v_p(a) < m$, on a $v_p(x^2) = 2s = v_p(a + yp^m) = \ell$. Ainsi, $\ell = 2s$, qui est pair. En simplifiant par p^ℓ, on obtient alors

$$x_0^2 = a_0 + yp^{m-\ell}.$$

Si $p > 2$, comme $m - \ell \geqslant 1$, on en déduit que a_0 est un carré modulo p. Si $p = 2$, on obtient également que a_0 est un carré modulo 4 si $m = \ell+2$, et un carré modulo 8 si $m \geqslant \ell+3$. Comme a_0 est impair, un calcul facile montre que cela équivaut à $a_0 \equiv 1$ [4] dans le premier cas, et à $a_0 \equiv 1$ [8] dans le second. Ainsi, les conditions énoncées dans le lemme sont nécessaires.

Si $p > 2$, ou si $p = 2$ et $m \geqslant \ell + 3$, le lemme 3.2 montre que ces conditions entraînent que a est un carré modulo p^k pour tout $k \geqslant 1$, donc modulo p^m.

Supposons que $p = 2$, $m = \ell + 1$ et que ℓ soit pair. Comme a_0 est impair, on a $a \equiv 2^\ell$ $[2^{\ell+1}]$, soit $a \equiv (2^{\frac{\ell}{2}})^2$ $[2^m]$. Ainsi, a est un carré modulo 2^m.

Supposons finalement que $p = 2$, $m = \ell+2$, que ℓ soit pair, et que $a_0 \equiv 1$ [4]. On a alors $a \equiv 2^\ell$ $[2^{\ell+2}]$, soit $a \equiv (2^{\frac{\ell}{2}})^2$ $[2^m]$, et a est un carré modulo 2^m. Cela achève la démonstration. □

Le corollaire suivant est contenu dans l'énoncé de la proposition précédente.

3.5. Corollaire. *Soit p un nombre premier, soit $a \in \mathbb{Z}$ un entier non nul et soit $m \geqslant 1$. Écrivons $a = p^\ell a_0$, avec $\ell \geqslant 0$ et $p \nmid a_0$.*

On suppose que $m > \ell$ si p est impair, et que $m \geqslant \ell + 3$ si $p = 2$. Alors, a est un carré modulo p^k pour tout $k \geqslant 1$ si, et seulement si, a est un carré modulo p^m.

On finit ces considérations en déterminant dans certains cas le nombre de solutions de l'équation $x^2 \equiv a$ $[p^m]$.

3.6. Lemme. *Soit p un nombre premier, soit $y \in \mathbb{Z}$ un entier tel que $p \nmid y$ et soit $m \geqslant 1$. Pour tout $x \in \mathbb{Z}$,*

(1) *on a $x^2 \equiv y^2$ $[2^m]$ $\Longrightarrow$ $x \equiv \pm y$ $[2^m]$ ou $x \equiv \pm y + 2^{m-1}$ $[2^m]$,*

(2) *si $p > 2$, on a $x^2 \equiv y^2$ $[p^m]$ $\Longrightarrow$ $x \equiv \pm y$ $[p^m]$.*

Démonstration

(1) Procédons par récurrence sur m. Si $m = 1$, 2 ou 3, cela résulte d'un simple calcul. Supposons que le résultat soit vrai au rang $m \geqslant 3$, et montrons-le au rang $m + 1$.

Si $x^2 \equiv y^2\ [2^{m+1}]$, alors $x^2 \equiv y^2\ [2^m]$. Par hypothèse de récurrence, il existe un entier $m \in \mathbb{Z}$ tel que

$$x = \pm y + 2^m k \quad \text{ou} \quad x = \pm y + 2^{m-1} + 2^m k.$$

Si $x = \pm y + 2^m k$, on a $x \equiv \pm y\ [2^{m+1}]$ si k est pair et $x \equiv \pm y + 2^m\ [2^{m+1}]$ si k est impair, et l'on a fini.

Supposons que $x = \pm y + 2^{m-1} + 2^m k = \pm y + 2^{m-1}(2k+1)$. Puisque $m \geqslant 3$, on a $2m - 2 \geqslant m + 1$, et en élevant au carré, on obtient

$$x^2 \equiv y^2 \pm 2^m y\ [2^{m+1}].$$

On en déduit alors que $2^m y \equiv 0\ [2^{m+1}]$, d'où une contradiction puisque y est impair. Cela achève la récurrence.

(2) On procède encore par récurrence sur m. Si $m = 1$, le résultat est vrai, car $\mathbb{Z}/p\mathbb{Z}$ est un anneau intègre. Supposons le résultat vrai au rang $m \geqslant 1$, et montrons-le au rang $m + 1$. Si $x^2 \equiv y^2\ [p^{m+1}]$, alors $x^2 \equiv y^2\ [p^m]$. Par hypothèse de récurrence, il existe un entier $k \in \mathbb{Z}$ tel que $x = \pm y + p^m k$. Comme $m \geqslant 1$, on a $2m \geqslant m + 1$. On a alors

$$x^2 \equiv y^2 \pm 2kyp^m\ [p^{m+1}],$$

d'où $2kyp^m \equiv 0\ [p^{m+1}]$. Par conséquent, on a $2ky \equiv 0\ [p]$. Comme p est impair et que $p \nmid y, 2y$ est inversible modulo p et l'on en déduit que $p \mid k$. Mais alors, on a $x \equiv \pm y\ [p^{m+1}]$ puisque $x = \pm y + p^m k$. Cela achève la récurrence, ainsi que la démonstration. □

Il reste donc à trouver une méthode pour déterminer à quelles conditions un entier $a \in \mathbb{Z}$ premier avec p est un carré modulo p. C'est l'objet de la section suivante.

4. Symboles de Legendre et de Jacobi

Dans tout ce paragraphe, p désignera un nombre premier, et l'on note $\mathbb{F}_p$ l'anneau $\mathbb{Z}/p\mathbb{Z}$. Si $x \in \mathbb{Z}$, sa classe modulo p sera notée $\overline{x}$ ou $[x]_p$, selon le contexte [6].

Commençons par déterminer l'ordre du groupe quotient $\mathbb{F}_p^\times / \mathbb{F}_p^{\times 2}$. On peut toujours supposer p impair, le cas $p = 2$ étant trivial (le groupe $\mathbb{F}_p^\times$ et son quotient étant réduits au neutre dans ce cas).

6. La seconde notation étant réservée au cas où l'on travaillerait dans différents anneaux quotients.

4.1. Lemme. *Pour tout nombre premier p impair, $\mathbb{F}_p^\times/\mathbb{F}_p^{\times 2}$ est un groupe d'ordre 2. Autrement dit, on a*

$$|\mathbb{F}_p^{\times 2}| = \frac{p-1}{2}.$$

Démonstration. L'application d'élévation au carré

$$\begin{aligned} \varphi : \mathbb{F}_p^\times &\longrightarrow \mathbb{F}_p^\times \\ \overline{x} &\longmapsto \overline{x}^2 \end{aligned}$$

est un morphisme de groupes abéliens, d'image ${\mathbb{F}_p}^{\times 2}$. Par ailleurs, $\mathbb{F}_p$ étant intègre puisque p est premier, on a

$$\overline{x}^2 = \overline{1} \iff (\overline{x}-\overline{1})(\overline{x}+\overline{1}) = \overline{0} \iff \overline{x} = \{\pm\overline{1}\}.$$

Ainsi, $\mathrm{Ker}(\varphi)$ est d'ordre 2. Le théorème de factorisation donne alors en particulier

$$|\mathbb{F}_p^{\times 2}| = \frac{|\mathbb{F}_p^\times|}{2},$$

ce qui implique aisément le résultat souhaité. □

Nous allons maintenant énoncer et démontrer un critère dû à Euler.

4.2. Lemme. *Soit p un nombre premier impair. Alors, pour tout $\overline{x} \in {\mathbb{F}_p}^\times$, on a $\overline{x}^{\frac{p-1}{2}} = \pm\overline{1}$. De plus, $\overline{x}^{\frac{p-1}{2}} = \overline{1}$ si, et seulement si, $\overline{x} \in \mathbb{F}_p^{\times 2}$.*

Démonstration. Soit $\overline{x} \in {\mathbb{F}_p}^\times$. D'après le théorème de Lagrange (ou le petit théorème de Fermat), on a

$$\left(\overline{x}^{\frac{p-1}{2}}\right)^2 = \overline{x}^{p-1} = \overline{1}.$$

Par intégrité de ${\mathbb{F}_p}^\times$, on en déduit que $\overline{x}^{\frac{p-1}{2}} = \pm\overline{1}$. Par ailleurs, si $\overline{x} = \overline{y}^2$, on a

$$\left(\overline{x}^{\frac{p-1}{2}}\right)^2 = \overline{y}^{p-1} = \overline{1}.$$

Puisque $\mathbb{F}_p$ est un corps, donc intègre, le polynôme $X^{\frac{p-1}{2}} - \overline{1} \in \mathbb{F}_p[X]$ possède au plus $\dfrac{p-1}{2}$ racines dans $\mathbb{F}_p$ d'après le lemme XV-6.21. Or, nous venons de constater que l'ensemble des racines de ce polynôme contient $\mathbb{F}_p^{\times 2}$, qui possède $\dfrac{p-1}{2}$ éléments par le lemme précédent. Ainsi, l'ensemble des racines de $X^{\frac{p-1}{2}} - \overline{1}$ est exactement $\mathbb{F}_p^{\times 2}$, ce qui est ce que l'on voulait démontrer. □

Pour tout p impair, on note $\mathscr{L}_p : {\mathbb{F}_p}^\times \longrightarrow \{\pm\overline{1}\}$ le morphisme de groupes

obtenu par composition du morphisme

$$\begin{aligned} {\mathbb{F}_p}^\times &\longrightarrow \{\pm\overline{1}\} \\ \overline{x} &\longmapsto \overline{x}^{\frac{p-1}{2}} \end{aligned}$$

et de l'isomorphisme évident $\{\pm\overline{1}\} \simeq \{\pm 1\}$.

Pour tout $\overline{x} \in {\mathbb{F}_p}^\times$, on a donc $\mathscr{L}_p(\overline{x}) = 1$ si, et seulement si, $\overline{x} \in {\mathbb{F}_p}^{\times 2}$.

Pour tout nombre premier p impair, et pour tout $a \in \mathbb{Z}$, on pose

$$\left(\frac{a}{p}\right) = \begin{cases} \mathscr{L}_p(\overline{a}) & \text{si } p \nmid a \\ 0 & \text{si } p \mid a. \end{cases}$$

Autrement dit, $\left(\frac{a}{p}\right)$ est l'unique représentant de $\overline{a}^{\frac{p-1}{2}}$ dans $\{0, 1, -1\}$ (et cela, même si $p \mid a$, puisque dans ce cas, $\overline{a} = \overline{0}$). Les propriétés suivantes sont alors immédiates d'après les définitions.

4.3. Lemme. *Soit p un nombre premier impair. Alors :*

(1) *pour tous $a, b \in \mathbb{Z}$ tels que $a \equiv b\ [p]$, on a $\left(\frac{a}{p}\right) = \left(\frac{b}{p}\right)$;*

(2) *pour tous $a \in \mathbb{Z}$, et tout $b \geqslant 1$ premier à p, on a $\left(\frac{ab^2}{p}\right) = \left(\frac{a}{p}\right)$;*

(3) *pour tous $a, b \in \mathbb{Z}$, on a $\left(\frac{ab}{p}\right) = \left(\frac{a}{p}\right)\left(\frac{b}{p}\right)$;*

(4) *pour tout $a \in \mathbb{Z}$, on a $\left(\frac{a}{p}\right) = 0$ si, et seulement si, $p \mid a$;*

(5) *pour tout $a \in \mathbb{Z}$ premier à p, on a $\left(\frac{a}{p}\right) = 1$ si, et seulement si, a est un carré modulo p.*

4.4. Définition. Pour tout p premier, l'application $\left(\frac{\cdot}{p}\right)$ s'appelle le *symbole de Legendre modulo p.*

Le lemme précédent montre en particulier que pour calculer la valeur du symbole $\left(\frac{a}{p}\right)$, il suffit de savoir calculer les symboles $\left(\frac{-1}{p}\right)$ et $\left(\frac{q}{p}\right)$ pour tout nombre premier q distinct de p. En guise d'échauffement, nous allons calculer les symboles de Legendre de -1 et 2.

4.5. Proposition. *Soit p un nombre premier impair.*

(1) *On a $\left(\frac{-1}{p}\right) = (-1)^{\frac{p-1}{2}}$.*

Autrement dit, on a

$$\left(\frac{-1}{p}\right) = \begin{cases} 1 & si \quad p \equiv 1 \ [4] \\ -1 & si \quad p \equiv 3 \ [4]. \end{cases}$$

(2) *On a* $\left(\frac{2}{p}\right) = (-1)^{\frac{p^2-1}{8}}$.

Autrement dit, on a

$$\left(\frac{2}{p}\right) = \begin{cases} 1 & si \quad p \equiv \pm 1 \ [8] \\ -1 & si \quad p \equiv \pm 3 \ [8]. \end{cases}$$

Démonstration. Le premier point découle simplement des définitions. Pour démontrer le second point, écrivons $p = 2s+1$, $s \geqslant 1$. Si $s = 2m$ ou $2m-1$, alors le plus grand entier positif impair inférieur à s est $2m-1$, tandis que le plus grand entier pair inférieur à s est $2(s-m)$. On a alors la série d'égalités suivantes :

$$\begin{aligned} s! &= \prod_{k=1}^{s} (-1)^k (-1)^k k \\ &= (-1)^{\frac{s(s+1)}{2}} \prod_{k=1}^{s} (-1)^k k \\ &= (-1)^{\frac{s(s+1)}{2}} \Big(\prod_{i=1}^{s-m} 2i \Big) \Big(\prod_{j=0}^{m-1} (-2j-1) \Big) \\ &= (-1)^{\frac{s(s+1)}{2}} 2^{s-m} (s-m)! \Big(\prod_{j=0}^{m-1} (-2j-1) \Big). \end{aligned}$$

Par définition de s, on a $-1 \equiv 2s \ [p]$. Ainsi, on a

$$\begin{aligned} \prod_{j=0}^{m-1} (-2j-1) &\equiv \prod_{j=0}^{m-1} (-2j+2s) \ [p] \\ &\equiv 2^m \prod_{j=0}^{m-1} (s-j) \ [p] \\ &\equiv 2^m \prod_{\ell=s-m+1}^{s} \ell \ [p]. \end{aligned}$$

Finalement, on obtient les congruences

$$s! \equiv (-1)^{\frac{s(s+1)}{2}} 2^{s-m} 2^m (s-m)! \prod_{\ell=s-m+1}^{s} \ell \equiv (-1)^{\frac{s(s+1)}{2}} 2^s s! \ [p].$$

Comme $s < p$ et que p est premier, $s!$ est inversible modulo p. En réarran-

geant les termes, on obtient donc

$$2^s \equiv (-1)^{\frac{s(s+1)}{2}} \ [p].$$

Or, $s = \dfrac{p-1}{2}$ et $\dfrac{s(s+1)}{2} = \dfrac{p-1}{2} \times \dfrac{p+1}{4} = \dfrac{p^2-1}{8}$. Ainsi, on a

$$\overline{2}^{\frac{p-1}{2}} = \overline{2}^s = (\overline{-1})^{\frac{p^2-1}{8}}.$$

On obtient alors l'égalité annoncée en utilisant la définition du symbole de Legendre. Un nombre premier impair étant nécessairement congru à ± 1, ± 3 modulo 8, la dernière partie résulte simplement des congruences

$$(\pm 3 + 8k)^2 \equiv 9 \ [16] \quad \text{et} \quad (\pm 1 + 8k)^2 \equiv 1 \ [16], \qquad \text{pour tout } k \in \mathbb{Z}.$$

Cela achève la démonstration. □

Nous allons maintenant établir la loi de réciprocité quadratique, qui nous permettra de calculer le symbole de Legendre, de manière (presque) algorithmique.

4.6. Théorème (Loi de réciprocité quadratique). *Soient p, q deux nombres premiers impairs distincts. Alors, on a*

$$\left(\frac{p}{q}\right)\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2} \times \frac{q-1}{2}}.$$

Autrement dit, on a

$$\left(\frac{q}{p}\right) = \begin{cases} -\left(\dfrac{p}{q}\right) & \text{si } p \equiv q \equiv 3 \ [4] \\ \left(\dfrac{p}{q}\right) & \text{sinon.} \end{cases}$$

Démonstration. La jolie démonstration peu connue que nous allons présenter ici est due à G. Rousseau [40]. Elle est basée sur les deux observations suivantes. Soit H le sous-groupe de $(\mathbb{Z}/p\mathbb{Z})^\times \times (\mathbb{Z}/q\mathbb{Z})^\times$ d'ordre 2 engendré par $([-1]_p, [-1]_q)$.

Fait No 1. Si $\varepsilon_1, \ldots, \varepsilon_4 \in \{\pm 1\}$ vérifient $([\varepsilon_1]_p, [\varepsilon_2]_q)H = ([\varepsilon_3]_p, [\varepsilon_4]_q)H$, alors

$$\varepsilon_1 \varepsilon_2 = \varepsilon_3 \varepsilon_4.$$

En effet, par hypothèse, il existe $\varepsilon \in \{\pm 1\}$ tel que

$$[\varepsilon_1]_p = [\varepsilon\varepsilon_3]_p \quad \text{et} \quad [\varepsilon_2]_q = [\varepsilon\varepsilon_4]_q.$$

Or, tout élément de $\mathbb{Z}/p\mathbb{Z}$ possède un unique représentant contenu dans l'intervalle $\left[\!\left[-\dfrac{p-1}{2}, \dfrac{p-1}{2}\right]\!\right]$.

Comme ε_1 et $\varepsilon\varepsilon_3$ sont égaux à ± 1, on en déduit que $\varepsilon_1 = \varepsilon\varepsilon_3$. On obtient de même que $\varepsilon_2 = \varepsilon\varepsilon_4$, d'où les égalités

$$\varepsilon_1\varepsilon_2 = \varepsilon^2\varepsilon_3\varepsilon_4 = \varepsilon_3\varepsilon_4.$$

Fait No 2. Si $([a_1]_p, [b_1]_q), \ldots, ([a_r]_p, [b_r]_q)$ et $([c_1]_p, [d_1]_q), \ldots, ([c_r]_p, [d_r]_q)$ sont deux systèmes complets de représentants des éléments du groupe quotient

$$G = \big((\mathbb{Z}/p\mathbb{Z})^\times \times (\mathbb{Z}/q\mathbb{Z})^\times\big)/H,$$

on a alors l'égalité

$$\Big(\Big[\prod_{k=1}^{r} a_k\Big]_p, \Big[\prod_{k=1}^{r} b_k\Big]_q\Big)H = \Big(\Big[\prod_{k=1}^{r} c_k\Big]_p, \Big[\prod_{k=1}^{r} d_k\Big]_q\Big)H.$$

En effet, chaque membre de cette égalité n'est rien d'autre que le produit de tous les éléments du groupe abélien G, éventuellement pris dans un ordre différent, ce qui n'a aucune importance puisque G est abélien.

Ces deux observations étant faites, passons à la démonstration proprement dite. Remarquons tout d'abord que l'ensemble

$$E = \Big\{([i]_p, [j]_q) \mid (i,j) \in [\![1, p-1]\!] \times \Big[\![1, \frac{q-1}{2}\Big]\!]\Big\}$$

est un système complet de représentants des éléments de G. En effet, tout élément de $\mathbb{Z}/q\mathbb{Z}$ admet un unique représentant dans $\Big[\!\Big[-\frac{q-1}{2}, \frac{q-1}{2}\Big]\!\Big]$. Quitte à multiplier par $([-1]_p, [-1]_q) \in H$, on en déduit que tout élément de G admet un représentant de la forme

$$([i]_p, [j]_q), \quad (i,j) \in [\![0, p-1]\!] \times \Big[\![0, \frac{q-1}{2}\Big]\!].$$

On remarque enfin que i et j ne peuvent être non nuls, puisque l'on travaille avec des classes inversibles, et donc que $([i]_p, [j]_q) \in E$.

L'unicité étant claire, on a le résultat souhaité. Le produit A de tous les classes modulo H des éléments de E est donc égal à

$$A = \Big([(p-1)!]_p^{\frac{q-1}{2}}, \ [(\frac{q-1}{2})!]_q^{p-1}\Big)H.$$

Remarquons maintenant que l'on a

$$\begin{aligned}\Big(\frac{q-1}{2}\Big)!^2 &\equiv \Big(\frac{q-1}{2}\Big)! \prod_{k=1}^{\frac{q-1}{2}} (k-q) \quad [q]\\ &\equiv (-1)^{\frac{q-1}{2}}\Big(\frac{q-1}{2}\Big)! \prod_{k=1}^{\frac{q-1}{2}} (q-k)\ [q]\\ &\equiv (-1)^{\frac{q-1}{2}}\Big(\frac{q-1}{2}\Big)! \prod_{k=\frac{q+1}{2}}^{q-1} k \quad [q]\\ &\equiv (-1)^{\frac{q-1}{2}}(q-1)! \quad [q].\end{aligned}$$

Ainsi, on a

$$A = \Big(\big[(p-1)!\big]_p^{\frac{q-1}{2}},\ \big[(-1)^{\frac{p-1}{2}\frac{q-1}{2}}\big]_q \big[(q-1)!\big]_q^{\frac{p-1}{2}}\Big)H.$$

Nous allons maintenant exhiber un autre système complet de représentants des éléments de G à l'aide du théorème chinois. L'isomorphisme

$$\begin{aligned}\mathbb{Z}/pq\mathbb{Z} &\longrightarrow \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}\\ [x]_{pq} &\longmapsto ([x]_p, [x]_q)\end{aligned}$$

se restreint en un isomorphisme de groupes

$$(\mathbb{Z}/pq\mathbb{Z})^\times \simeq (\mathbb{Z}/p\mathbb{Z})^\times \times (\mathbb{Z}/q\mathbb{Z})^\times.$$

De plus, l'image du sous-groupe engendré par $[-1]_{pq}$ via cet isomorphisme est le sous-groupe H. Or, l'ensemble

$$\Big\{[k]_{pq} \in \mathbb{Z}/pq\mathbb{Z} \;\Big|\; k \in \Big[\!\Big[1, \frac{pq-1}{2}\Big]\!\Big],\ \operatorname{pgcd}(k, pq) = 1\Big\}$$

est un système complet de représentants des éléments du groupe quotient

$$(\mathbb{Z}/pq\mathbb{Z})^\times / \langle [-1]_{pq} \rangle.$$

L'ensemble

$$E' = \{([k]_p, [k]_q) \in \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z} \mid k \in I\},$$

où $I = \Big\{k \in \Big[\!\Big[1, \frac{pq-1}{2}\Big]\!\Big] \mid \operatorname{pgcd}(k, pq) = 1\Big\}$, est alors un système complet de représentants des éléments de G. Comme précédemment, nous allons calculer le produit B des éléments de E'.

Soit F l'ensemble des éléments de $\Big[\!\Big[1, \frac{pq-1}{2}\Big]\!\Big]$ qui sont premiers avec p. Un élément de F est soit premier avec pq, soit un multiple de q. Remarquons que pour tout entier $j \geqslant 0$, on a

$$1 \leqslant jq \leqslant \frac{pq-1}{2} \quad \Longleftrightarrow \quad 1 \leqslant j \leqslant \frac{p-1}{2}.$$

On a donc

$$\prod_{k\in F}[k]_p = \Big(\prod_{k\in I}[k]_p\Big)\Big(\prod_{j=1}^{(p-1)/2}[jq]_p\Big) = \Big(\prod_{k\in I}[k]_p\Big)\big[(\tfrac{p-1}{2})!\big]_p\,[q]_p^{\frac{p-1}{2}}.$$

Par ailleurs, un élément de F est de la forme

$$jp+i, \quad i\in[\![1,p-1]\!].$$

On voit aisément que si $j \geqslant \dfrac{q+1}{2}$, alors $jp+i > \dfrac{pq-1}{2}$. Par conséquent, on a $j \leqslant \dfrac{q-1}{2}$.

Si $j \in \Big[\!\Big[1, \dfrac{q-1}{2}-1\Big]\!\Big]$, on vérifie facilement que $jp+i \in \Big[\!\Big[1, \dfrac{pq-1}{2}\Big]\!\Big]$, pour tout $i\in[\![1,p-1]\!]$.

Si $j = \dfrac{q-1}{2}$, on constate que $jp+i \in \Big[\!\Big[1, \dfrac{pq-1}{2}\Big]\!\Big]$ si, et seulement si, on a $i \in \Big[\!\Big[1, \dfrac{p-1}{2}\Big]\!\Big]$.

On déduit de tout cela les égalités suivantes :

$$\begin{aligned}\prod_{k\in F}[k]_p &= \Big(\prod_{j=0}^{(q-3)/2}\prod_{i=1}^{p-1}[jp+i]_p\Big)\Big(\prod_{i=1}^{(p-1)/2}\big[\tfrac{q-1}{2}p+i\big]_p\Big)\\ &= \big[(p-1)!\big]_p^{\frac{q-1}{2}}\big[(\tfrac{p-1}{2})!\big]_p.\end{aligned}$$

En comparant avec le résultat ci-dessus, puis en utilisant l'inversibilité de $\big[(\dfrac{p-1}{2})!\big]_p$ modulo p, et le fait que $[q]_p^{\frac{p-1}{2}}$ est son propre inverse (puisqu'égal à $[\pm 1]_p$), on en déduit l'égalité

$$\prod_{k\in I}[k]_p = [(p-1)!]_p^{\frac{q-1}{2}}\,[q]_p^{\frac{p-1}{2}}.$$

En échangeant les rôles de p et q, on obtient aussi

$$\prod_{k\in I}[k]_q = [(q-1)!]_q^{\frac{p-1}{2}}\,[p]_q^{\frac{q-1}{2}}.$$

Ainsi, on a

$$B = \Big([(p-1)!]_p^{\frac{q-1}{2}}\,[q]_p^{\frac{p-1}{2}},\ [(q-1)!]_q^{\frac{p-1}{2}}\,[p]_q^{\frac{q-1}{2}}\Big)H.$$

D'après la seconde observation faite en début de démonstration, on a l'égalité $A = B$.

La simplification par $\left([(p-1)!]_p^{\frac{q-1}{2}}, [(q-1)!]_q^{\frac{p-1}{2}}\right)$ dans cette égalité donne

$$\left([1]_p, [(-1)^{\frac{p-1}{2}\frac{q-1}{2}}]_q\right)H = \left([q]_p^{\frac{p-1}{2}}, [p]_q^{\frac{q-1}{2}}\right)H.$$

D'après la définition du symbole de Legendre, cela se récrit

$$\left([1]_p, \left[(-1)^{\frac{p-1}{2}\times\frac{q-1}{2}}\right]_q\right)H = \left(\left[\left(\frac{q}{p}\right)\right]_p, \left[\left(\frac{p}{q}\right)\right]_q\right)H.$$

La première observation faite précédemment nous donne alors l'égalité désirée. Le dernier point provient d'un simple calcul. Cela achève la démonstration. □

Le lemme 4.3, la proposition 4.5 et le théorème de réciprocité quadratique donnent un « algorithme » pour calculer $\left(\frac{a}{p}\right)$.

Algorithme

(1) On réduit a modulo p, et l'on remplace a par l'unique représentant de $\overline{a}$ dans l'intervalle $\left[\!\left[-\frac{p-1}{2}, \frac{p-1}{2}\right]\!\right]$.

(2) Si $a = 0$, on a fini.

On suppose que $a \neq 0$, i.e. que a est premier avec p.

(3) On écrit $a = \varepsilon\left(\prod_{q\in E} q\right)b^2$, avec $\varepsilon = \pm 1$, et où E est un ensemble fini de nombres premiers. On a alors

$$\left(\frac{a}{p}\right) = \left(\frac{\varepsilon}{p}\right)\prod_{q\in E}\left(\frac{q}{p}\right).$$

Si $q = 2$, on calcule $\left(\frac{q}{p}\right)$ en utilisant la proposition 4.5. Pour tout nombre premier $q \in E$ impair, on écrit

$$\left(\frac{q}{p}\right) = (-1)^{\frac{p-1}{2}\times\frac{q-1}{2}}\left(\frac{p}{q}\right),$$

puis on amorce le calcul de $\left(\frac{p}{q}\right)$ en repartant à l'étape (1), jusqu'à finalement être réduit à calculer des symboles $\left(\frac{c}{\ell}\right)$, où ℓ est premier impair et $c \in \{1, -1, 2\}$.

(4) On fait le produit des symboles calculés au fur et à mesure.

4.7. Exemple. Calculons $\left(\frac{103}{173}\right)$. On a $103 \equiv -70\ [173]$, donc

$$\left(\frac{103}{173}\right) = \left(\frac{-1}{173}\right)\left(\frac{2}{173}\right)\left(\frac{5}{173}\right)\left(\frac{7}{173}\right).$$

En utilisant tous les résultats précédents, on obtient

$$\left(\frac{103}{173}\right) = (1)(-1)\left(\frac{173}{5}\right)\left(\frac{173}{7}\right),$$

soit encore

$$\left(\frac{103}{173}\right) = -\left(\frac{-2}{5}\right)\left(\frac{-2}{7}\right).$$

Ainsi, on obtient

$$\left(\frac{103}{173}\right) = -\left(\frac{-1}{5}\right)\left(\frac{-1}{7}\right)\left(\frac{2}{5}\right)\left(\frac{2}{7}\right) = -(1)(-1)(-1)^3(-1)^6 = -1.$$

Autrement dit, 103 n'est pas un carré modulo 173.

Remarquons que, 103 étant premier, on aurait aussi pu utiliser la loi de réciprocité quadratique immédiatement, mais l'avantage de la première étape de l'algorithme est de faire intervenir des nombres premiers plus petits, et de raccourcir ainsi le calcul.

Le terme « algorithme » est plutôt galvaudé ici, puisque pour le mettre en œuvre, il faut avoir un algorithme de factorisation d'entiers efficace, ce dont on ne dispose pas pour les grands entiers à l'heure actuelle.

Pour éviter cet écueil, nous allons introduire un nouveau symbole, qui généralise le symbole de Legendre.

4.8. Définition. Soit $N \geqslant 2$ un entier **impair**, et soit $N = p_1^{m_1} \times \cdots \times p_r^{m_r}$ sa décomposition en facteurs premiers. Pour tout $a \in \mathbb{Z}$, on pose

$$\left(\frac{a}{N}\right) = \left(\frac{a}{p_1}\right)^{m_1} \cdots \left(\frac{a}{p_r}\right)^{m_r}.$$

L'application $\left(\frac{\cdot}{N}\right)$ s'appelle le *symbole de Jacobi modulo* N.

4.9. Remarque. Si N est un nombre premier impair, $\left(\frac{\cdot}{N}\right)$ n'est donc rien d'autre que le symbole de Legendre modulo N.

On a alors les propriétés suivantes, qui découlent immédiatement de la définition et des propriétés du symbole de Legendre.

4.10. Lemme

Le symbole de Jacobi possède les propriétés suivantes :

(1) *pour tous* $a, b \in \mathbb{Z}$, *et tout* $N \geqslant 2$ *impair, on a* $\left(\frac{ab}{N}\right) = \left(\frac{a}{N}\right)\left(\frac{b}{N}\right)$;

(2) *pour tout* $N \geqslant 2$ *impair, tout* $a \in \mathbb{Z}$, *et tout entier* $b \geqslant 1$ *premier à* N, *on a* $\left(\frac{ab^2}{N}\right) = \left(\frac{a}{N}\right)$;

(3) *pour tout* $N \geqslant 2$ *impair, tout* $D \geqslant 1$ *impair, et tout* $a \in \mathbb{Z}$, *on a l'égalité* $\left(\frac{a}{ND^2}\right) = \left(\frac{a}{N}\right)$;

(4) *pour tous* M, $N \geqslant 2$ *impairs et premiers entre eux, et tout entier* $a \in \mathbb{Z}$, *on a* $\left(\frac{a}{MN}\right) = \left(\frac{a}{M}\right)\left(\frac{a}{N}\right)$;

(5) *pour tout* $N \geqslant 2$ *impair, et tous entiers* $a, b \in \mathbb{Z}$ *tels que* $a \equiv b\ [N]$, *on a* $\left(\frac{a}{N}\right) = \left(\frac{b}{N}\right)$;

(6) *pour tout* $N \geqslant 2$ *impair, et tout* $a \in \mathbb{Z}$, *on a* $\left(\frac{a}{N}\right) \neq 0$ *si, et seulement si,* a *et* N *sont premiers entre eux.*

4.11. Remarque. Soit $a \in \mathbb{Z}$ et $N \geqslant 2$ un entier impair. Si $\left(\frac{a}{N}\right) = -1$, alors a n'est pas un carré modulo N. En effet, par définition, il existe alors au moins un diviseur premier p de N tel que $\left(\frac{a}{p}\right) = -1$. En particulier, a n'est pas un carré modulo p, et donc a fortiori modulo N.

En revanche, si $\left(\frac{a}{N}\right) = 1$, rien n'assure que a soit un carré modulo N. Par exemple, on a

$$\left(\frac{-1}{21}\right) = \left(\frac{-1}{3}\right)\left(\frac{-1}{7}\right) = (-1)(-1) = 1,$$

mais -1 n'est pas un carré modulo 21, puisque par exemple ce n'est pas un carré modulo 3.

Ainsi, contrairement au symbole de Legendre, le symbole de Jacobi ne caractérise pas les carrés modulo N.

On continue par un lemme technique.

4.12. Lemme

Soient M, $N \geqslant 1$ *deux entiers impairs. Alors :*

(1) *on a* $\frac{M-1}{2} + \frac{N-1}{2} \equiv \frac{MN-1}{2}\ [2]$;

(2) *on a* $\dfrac{M^2-1}{8}+\dfrac{N^2-1}{8}\equiv\dfrac{(MN)^2-1}{8}$ [2].

Démonstration. Commençons par démontrer (1).

Si $M\equiv N$ [4], alors $MN\equiv 1$ [4]. Ainsi, $\dfrac{M-1}{2}$, $\dfrac{N-1}{2}$ sont de même parité et $\dfrac{MN-1}{2}$ est pair.

Si $M\equiv 3$ [4] et $N\equiv 1$ [4] ou inversement, alors $MN\equiv 3$ [4]. Mais alors, les entiers $\dfrac{M-1}{2}$ et $\dfrac{N-1}{2}$ sont de parités différentes, et $\dfrac{MN-1}{2}$ est impair.

Dans tous les cas, la congruence désirée est vérifiée.

Démontrons (2).

Remarquons tout d'abord que si $a\in\mathbb{Z}$ est impair, alors a est congru à ± 1, ou ± 3 modulo 8, et que dans ce cas, a^2 est congru à 1 ou 9 modulo 16. De plus, dans tous les cas, on a $a^4\equiv 1$ [16].

Si $M^2\equiv N^2$ [16], alors $M^2N^2\equiv M^4\equiv 1$ [16]. Ainsi, $\dfrac{M^2-1}{8}$, $\dfrac{N^2-1}{8}$ sont de même parité et $\dfrac{(MN)^2-1}{8}$ est pair.

Si $M^2\equiv 1$ [16] et $N^2\equiv 9$ [16] ou inversement, alors $M^2N^2\equiv 9$ [16]. Mais alors, les entiers $\dfrac{M^2-1}{8}$ et $\dfrac{N^2-1}{8}$ sont de parités différentes, et ainsi $\dfrac{(MN)^2-1}{8}$ est impair.

Là encore, dans tous les cas, la congruence désirée est vérifiée. □

On a maintenant le théorème suivant, qui généralise les propriétés calculatoires du symbole de Legendre.

4.13. Théorème. *On a les identités suivantes :*

(1) *pour tout* $N\geqslant 2$ *impair, on a* $\left(\dfrac{-1}{N}\right)=(-1)^{\frac{N-1}{2}}$. *Autrement dit,*

$$\left(\frac{-1}{N}\right)=\begin{cases}1 & si \quad N\equiv 1\ [4]\\ -1 & si \quad N\equiv 3\ [4],\end{cases}$$

(2) *pour tout* $N\geqslant 2$ *impair, on a* $\left(\dfrac{2}{N}\right)=(-1)^{\frac{N^2-1}{8}}$. *Autrement dit,*

$$\left(\frac{2}{N}\right)=\begin{cases}1 & si \quad N\equiv \pm 1\ [8]\\ -1 & si \quad N\equiv \pm 3\ [8],\end{cases}$$

(3) *pour tous entiers M, $N \geqslant 2$ impairs premiers entre eux, on a*

$$\left(\frac{M}{N}\right)\left(\frac{N}{M}\right) = (-1)^{\frac{M-1}{2}\times\frac{N-1}{2}}.$$

Autrement dit,

$$\left(\frac{N}{M}\right) = \begin{cases} -\left(\dfrac{M}{N}\right) & \text{si } M \equiv N \equiv 3\ [4] \\ \left(\dfrac{M}{N}\right) & \text{sinon.} \end{cases}$$

Démonstration. Soit $N = p_1^{m_1} \times \cdots \times p_r^{m_r}$ la décomposition en facteurs premiers de N, chaque p_i étant impair. On a alors

$$\left(\frac{-1}{N}\right) = \left(\frac{-1}{p_1}\right)^{m_1} \cdots \left(\frac{-1}{p_r}\right)^{m_r} = (-1)^{\sum_{i=1}^{r} m_i \frac{p_i-1}{2}}.$$

Une utilisation répétée du lemme 4.12 (1) montre que

$$\sum_{i=1}^{r} m_i \frac{p_i - 1}{2} \equiv \frac{p_1^{m_1} \times \cdots \times p_r^{m_r} - 1}{2} \equiv \frac{N-1}{2}\ [2].$$

Cela démontre alors le point (1). Le point (2) se montre de la même manière, mais cette fois en utilisant le lemme 4.12 (2).

Gardons les notations du lemme, et soit $M = q_1^{n_1} \times \cdots \times q_s^{n_s}$ la décomposition en facteurs premiers de M, chaque q_j étant impair, et distinct des p_i. On a alors

$$\left(\frac{M}{N}\right) = \prod_{i=1}^{r}\prod_{j=1}^{s}\left(\frac{q_j}{p_i}\right)^{m_i n_j},$$

et de même

$$\left(\frac{N}{M}\right) = \prod_{i=1}^{r}\prod_{j=1}^{s}\left(\frac{p_i}{q_j}\right)^{m_i n_j}.$$

D'après la loi de réciprocité quadratique, on obtient alors

$$\left(\frac{M}{N}\right)\left(\frac{N}{M}\right) = (-1)^{\sum_{i=1}^{r}\sum_{j=1}^{s} m_i n_j \frac{p_i-1}{2}\times\frac{q_j-1}{2}}.$$

Or, on a

$$\sum_{i=1}^{r}\sum_{j=1}^{s} m_i n_j \frac{p_i-1}{2}\times\frac{q_j-1}{2} = \Big(\sum_{i=1}^{r} m_i \frac{p_i-1}{2}\Big)\Big(\sum_{j=1}^{s} n_j \frac{q_j-1}{2}\Big).$$

Comme on l'a déjà remarqué en début de démonstration, le premier facteur est congru à $\dfrac{N-1}{2}$ modulo 2, et le second est congru à $\dfrac{M-1}{2}$ modulo 2.

Ainsi,

$$\sum_{i=1}^{r}\sum_{j=1}^{s} m_i n_j \frac{p_i-1}{2}\times\frac{q_j-1}{2} \equiv \frac{M-1}{2}\times\frac{N-1}{2} \quad [2],$$

ce qui achève la démonstration. □

Ce théorème permet de calculer un symbole de Jacobi, et donc un symbole de Legendre de manière totalement algorithmique.

Notation. Si $c \in \mathbb{Z}$ et $n \geqslant 2$ est un entier impair, on note c_n l'unique représentant de la classe de c modulo n dans l'intervalle $\left[\!\left[-\frac{n-1}{2}, \frac{n-1}{2}\right]\!\right]$.

Algorithme[7]**.** On se donne un entier impair $N \geqslant 2$ et $a \in \mathbb{Z}$.

(1) En utilisant l'algorithme d'Euclide (non étendu), on vérifie si a et N sont premiers entre eux. Si ce n'est pas le cas, on a $\left(\frac{a}{N}\right) = 0$, et il n'a rien à faire.

On suppose donc maintenant que a et N sont premiers entre eux. On déclare trois variables J, b, $M \in \mathbb{Z}$, que l'on initialise comme suit :

$$J = 1, \quad b = a, \quad M = N.$$

(2) Faire $b \longleftarrow b_M$.

(3) Décomposer b sous la forme $b = \varepsilon 2^m s$, où $s \geqslant 1$ est impair, $m \geqslant 0$ et où $\varepsilon \in \{\pm 1\}$.

(4) Faire $J \longleftarrow \left(\frac{\varepsilon}{M}\right) J$ si m est pair, et faire $J \longleftarrow \left(\frac{\varepsilon}{M}\right)\left(\frac{2}{M}\right) J$ si m est impair.

(5) Si $s = 1$, afficher J. Fin de l'algorithme. Si $s \geqslant 3$, passer à l'étape (6).

(6) Si $M \equiv s \equiv 3$ [4], faire $J \longleftarrow -J$.

(7) Faire $b \longleftarrow M$ et $M \longleftarrow s$. Repartir à l'étape (2).

L'algorithme s'arrête nécessairement puisque l'entier s défini à l'étape (3) est $< M$. Ainsi, au vu de l'étape (7), la suite des modulo M successifs est strictement décroissante. On aura donc $s = 1$ au bout d'un nombre fini d'étapes.

4.14. Exemple. Calculons $\left(\frac{13898}{8911}\right)$. On a

$$\left(\frac{13898}{8911}\right) = \left(\frac{-3924}{8911}\right) = \left(\frac{-1}{8911}\right)\left(\frac{2}{8911}\right)^2\left(\frac{981}{8911}\right).$$

7. Un vrai, cette fois !

On a ainsi

$$\left(\frac{13898}{8911}\right) = -\left(\frac{981}{8911}\right) = -\left(\frac{8911}{981}\right) = -\left(\frac{82}{981}\right),$$

d'où

$$\left(\frac{13898}{8911}\right) = -\left(\frac{2}{981}\right)\left(\frac{41}{981}\right) = \left(\frac{41}{981}\right) = \left(\frac{981}{41}\right).$$

On a alors

$$\left(\frac{13898}{8911}\right) = \left(\frac{981}{41}\right) = \left(\frac{-3}{41}\right) = \left(\frac{-1}{41}\right)\left(\frac{3}{41}\right) = \left(\frac{3}{41}\right).$$

Enfin, on obtient

$$\left(\frac{13898}{8911}\right) = \left(\frac{41}{3}\right) = \left(\frac{-1}{3}\right) = -1.$$

4.15. Remarque. Si $a \in \mathbb{Z}$ est un carré modulo p, il existe un algorithme, appelé algorithme de Tonelli-Shanks, qui permet de calculer une racine carrée de a modulo p. Le lecteur pourra consulter [19] avec profit.

5. Exercices

E1. Soit $n \geqslant 1$ et soit $d \mid n$ un diviseur positif de n. Montrer que le morphisme de groupes évident

$$(\mathbb{Z}/n\mathbb{Z})^\times \longrightarrow (\mathbb{Z}/d\mathbb{Z})^\times$$

est surjectif.

Indication. Considérer le produit t des diviseurs premiers de n ne divisant pas d, et remarquer que d et t sont premiers entre eux.

E2. Soit $n \geqslant 1$. Pour tous $\overline{a}, \overline{b} \in \mathbb{Z}/n\mathbb{Z}$, montrer que $(\overline{a}) = (\overline{b})$ si, et seulement si, il existe $\overline{u} \in (\mathbb{Z}/n\mathbb{Z})^\times$ tel que $\overline{b} = \overline{u}\,\overline{a}$.

Indication. Se ramener au cas où n est une puissance non triviale d'un nombre premier p. Dans ce cas, écrire $a = a'p^k$, où $p \nmid a'$ et $k \in [\![0, n]\!]$.

E3. Soit $n \geqslant 2$ et soit $n = p_1^{m_1} \times \cdots \times p_r^{m_r}$ sa décomposition en facteurs premiers. Montrer que l'on a un isomorphisme de groupes

$$\mathrm{Nil}(\mathbb{Z}/n\mathbb{Z}) \simeq \mathbb{Z}/p_1^{m_1-1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_r^{m_r-1}\mathbb{Z}.$$

Indication. On pourra démontrer que pour tout nombre premier p, et tout $m \geqslant 1$, on a

$$\mathrm{Nil}(\mathbb{Z}/p^m\mathbb{Z}) = p\mathbb{Z}/p^m\mathbb{Z}.$$

E4. Déterminer tous les entiers $n \geqslant 2$ tels que $(\mathbb{Z}/n\mathbb{Z})^\times$ est cyclique.

E5. Déterminer les diviseurs de zéro, les éléments nilpotents et les éléments idempotents de l'anneau $\mathbb{Z}/60\mathbb{Z}$.

E6. Soit p un nombre premier impair, et soit $m \geqslant 1$.

a. Montrer que pour tout $k \in [\![0, m]\!]$, le morphisme de groupes

$$\mathbb{Z}/p^{m-k}\mathbb{Z} \longrightarrow \mathbb{Z}/p^m\mathbb{Z}$$
$$[x]_{p^{m-k}} \longmapsto [p^k x]_{p^m}$$

est bien défini, et induit un isomorphisme de groupes

$$(\mathbb{Z}/p^{m-k}\mathbb{Z})^{\times 2} \simeq p^k(\mathbb{Z}/p^m\mathbb{Z})^{\times 2}.$$

b. En déduire que pour tout $k \in [\![0, m]\!]$, on a

$$|p^k(\mathbb{Z}/p^m\mathbb{Z})^{\times 2}| = p^{m-k-1}\,\frac{p-1}{2}.$$

c. Montrer tout $[a]_{p^m}$ non nul s'écrit sous la forme

$$[a]_{p^m} = [p^k]_{p^m}[u]_{p^m},$$

avec $k \in [\![0, m-1]\!]$ et $p \nmid u$. Montrer également que l'entier k est unique, et que u est unique à un multiple de p^{m-k} près.

d. En déduire que si $m = 2s$, $|(\mathbb{Z}/p^m\mathbb{Z})^2|$ est égal à

$$1 + \frac{p-1}{2}(p + p^3 + \cdots + p^{2s-1}),$$

et que si $m = 2s+1$, il est égal à

$$1 + \frac{p-1}{2}(1 + p^2 + \cdots + p^{2s}).$$

Les exercices suivants nécessitent d'avoir étudié le paragraphe §4 de ce chapitre.

E7. Pour tout nombre premier impair p, calculer $\left(\frac{3}{p}\right)$.

E8. Calculer $\left(\frac{24680}{13579}\right)$.

E9. Soit p un nombre premier impair.

a. Vérifier qu'il existe un unique morphisme de groupes non trivial

$$\varphi : {\mathbb{F}_p}^{\times} \longrightarrow \{\pm 1\}.$$

Calculer φ en fonction de $\left(\frac{-}{p}\right)$.

Pour tout $\overline{a} \in {\mathbb{F}_p}^\times$, on note $\sigma_{\overline{a}} : {\mathbb{F}_p}^\times \longrightarrow {\mathbb{F}_p}^\times$ l'application de multiplication par $\overline{a}$.

b. Vérifier que $\sigma_{\overline{a}} \in \mathfrak{S}({\mathbb{F}_p}^\times)$, et montrer que si $\overline{a}_0$ est un générateur de ${\mathbb{F}_p}^\times$, on a $\varepsilon(\sigma_{\overline{a}_0}) = -1$.

c. En déduire que pour tout entier $a \in \mathbb{Z}$ premier à p, on a l'égalité

$$\varepsilon(\sigma_{\overline{a}}) = \left(\frac{a}{p}\right).$$

d. En utilisant l'exercice VIII-26, en déduire que pour tout entier $a \in \mathbb{Z}$ premier à p, on a

$$\left(\frac{a}{p}\right) = (-1)^{(o(\overline{a})-1)\frac{p-1}{o(\overline{a})}},$$

puis que a n'est pas un carré modulo p si, et seulement si, $\langle\overline{a}\rangle$ contient le 2-Sylow de ${\mathbb{F}_p}^\times$.

E10. Soit $D \in \mathbb{Z} \setminus \{0\}$ tel que $D \equiv 0, 1\ [4]$.

Montrer que pour tous entiers m, n **impairs** tels que $D \mid (m-n)$, on a

$$\left(\frac{D}{m}\right) = \left(\frac{D}{n}\right).$$

Chapitre XVIII

Divisibilité dans les anneaux

Un anneau pour les factoriser tous, un anneau pour les plonger dans une descente infinie et enfin démontrer cette conjecture infernale ![1]

On s'intéresse dans ce chapitre à la factorisation dans un anneau intègre. On veut en particulier savoir si le théorème fondamental de l'arithmétique pour les entiers (cf. théorème II-4.4) se généralise. Pour cela, on doit tout d'abord trouver des substituts à la notion de divisibilité, ainsi qu'à la notion de nombre premier.

1. Divisibilité, éléments irréductibles et premiers

1.1. Définition. Soit A un anneau commutatif, et soient a, $b \in A$. On dit que a *divise* b s'il existe $x \in A$ tel que $b = ax$, ou de manière équivalente, si l'on a $(b) \subset (a)$. On le note $a \mid b$.

On dit aussi que a est un *diviseur* de b, et que b est un *multiple* de a.

1.2. Remarques

(1) Cette définition soulève une petite difficulté lorsque $b = 0$. En effet, tout élément $a \in A$ est un diviseur de 0 au sens précédent (puisque l'on peut toujours écrire $0 = a\,0$), mais pas nécessairement un diviseur de zéro au sens de la définition XV-2.6.

1. Sauron de Mordor, alors qu'il était en train d'essayer de démontrer en vain le grand théorème de Fermat en factorisant en produits d'éléments irréductibles dans $\mathbb{Z}[\zeta_n]$.

(2) La notion de divisibilité est relative à un anneau ambiant fixé (cf. les points (2) et (3) de la série d'exemples qui suit). Il est donc important que l'élément a de la définition soit un élément de l'anneau A dans lequel on travaille.

1.3. Exemples

(1) Dans un anneau non trivial, 0 ne divise aucun élément non nul.

(2) Dans $\mathbb{Z}$, l'entier 2 ne divise pas 3. Sinon, on aurait $3 = 2m$ pour $m \in \mathbb{Z}$, ce qui est faux, puisque 3 est impair.

(3) On a $2 \mid 3$ dans $\mathbb{Q}$ puisque $3 = 2\,\frac{3}{2}$.

(4) Dans $\mathbb{Z}[i]$, 2 ne divise pas $1+i$. Sinon, on aurait $1+i = 2z,\ z \in \mathbb{Z}[i]$. En écrivant $z = a + bi,\ a, b \in \mathbb{Z}$, on obtient $1 + i = 2a + 2bi$, et en particulier $1 = 2a$ dans $\mathbb{Z}$, ce qui est impossible.

(5) On a $1 + i \mid 2$ dans $\mathbb{Z}[i]$ puisque $2 = (1+i)(1-i)$, et $1 - i \in \mathbb{Z}[i]$.

1.4. Définition. On dit que deux éléments $a, b \in A$ sont *associés* si $a \mid b$ et $b \mid a$, autrement dit si $(a) = (b)$. Si l'anneau A est intègre, cela revient à dire que $b = au,\ u \in A^\times$, d'après le lemme XV-4.14 (3).

La relation « être associés » est une relation d'équivalence.

1.5. Remarque. Si A est un anneau commutatif quelconque, alors deux éléments égaux à un inversible près sont clairement associés. En revanche, si l'anneau A n'est pas intègre, deux éléments peuvent être associés sans être égaux à un inversible près (cf. exercice 19).

Maintenant que l'on a la notion de divisibilité appropriée, on peut introduire la notion d'éléments premiers entre eux.

1.6. Définition. On dit que deux éléments a et b de A sont *premiers entre eux* si pour tout $d \in A$, on a

$$d \mid a \ \text{ et } \ d \mid b \quad \Longrightarrow \quad d \in A^\times.$$

Autrement dit, deux éléments sont premiers entre eux si leurs seuls diviseurs communs sont les éléments inversibles.

1.7. Exemple. Deux éléments $a, b \in A$ étrangers sont premiers entre eux.

En effet, supposons que a et b soient étrangers. Il existe donc $u, v \in A$ tels que $ua + bv = 1$. Soit $d \in A$ tels que $d \mid a$ et $d \mid b$. Alors, $d \mid (ua + vb)$, c'est-à-dire $d \mid 1$. La définition entraîne alors immédiatement que $d \in A^\times$.

1.8. Remarque. On peut démontrer que la réciproque est fausse en général (cf. exercice 2).

On peut aussi généraliser la notion de pgcd et de ppcm.

1.9. Définition. Soit A un anneau, et soient $a, b \in A$. On dit que $d \in A$ est un *pgcd* de a et b si d vérifie les propriétés suivantes :

(1) on a $d \mid a$ et $d \mid b$;

(2) pour tout $c \in A$ tel que $c \mid a$ et $c \mid b$, on a $c \mid d$.

On dit que $m \in A$ est un *ppcm* de a et b si m vérifie les propriétés suivantes :

(1) on a $a \mid m$ et $b \mid m$;

(2) pour tout $c \in A$ tel que $a \mid c$ et $b \mid c$, on a $m \mid c$.

Attention ! En général, un pgcd et un ppcm n'existent pas nécessairement. Par exemple, si $A = \mathbb{Z}[i\sqrt{5}]$, on peut montrer que 9 et $3(2+i\sqrt{5})$ n'ont pas de pgcd, et que 3 et $2+i\sqrt{5}$ n'ont pas de ppcm (cf. exercice 10).

1.10. Remarques

(1) Si $a = 0$, alors un pgcd de a et b est b. En effet, b divise à la fois 0 et b. Si maintenant $c \in A$ divise 0 et b, il divise b.
En particulier, si $a = b = 0$, un pgcd de a et b est 0.

(2) Si $(a, b) \neq (0, 0)$, alors un pgcd de a et b (lorsqu'il existe) est non nul. En effet, si par exemple $a \neq 0$, on a $a = da'$, $a' \in A$. Comme $a \neq 0$, on a nécessairement $d \neq 0$.

(3) Si $a = 0$ ou $b = 0$, alors 0 est un ppcm de a et b. En effet, 0 est un multiple commun à a et b. De plus, si c est un multiple commun à a et b, alors $c = 0$. Dans ce cas, c est bien multiple de 0.

(4) Si a et b sont non nuls, un ppcm m de a et b (s'il existe) est non nul dès que A est intègre, puisque m divise ab, qui est non nul par intégrité.

Le pgcd et le ppcm, lorsqu'ils existent, sont essentiellement uniques.

1.11. Lemme. *Soit A un anneau. Alors, un pgcd et un ppcm de deux éléments, lorsqu'ils existent, sont uniques à association près. Autrement dit, les idéaux engendrés par un pgcd et un ppcm respectivement sont indépendants du pgcd et du ppcm choisis.*

En particulier, si A est intègre, deux pgcd (resp. deux ppcm) sont égaux à multiplication par un élément inversible près.

Démonstration. On le montre pour le pgcd, le cas du ppcm étant similaire. Supposons que d et d' soient deux pgcd de a et b. Alors, puisque d' est un pgcd de a et b et que d est un diviseur commun à a et b, on a $d \mid d'$. De même, on a $d' \mid d$. Autrement dit, d et d' sont associés, ce qui revient à dire que $(d) = (d')$. Lorsque l'anneau A est intègre, cela équivaut à dire qu'il existe $u \in A^\times$ tel que $d = ud'$ (cf. définition 1.4). □

1.12. Remarque. Soient $a, b \in A$. Alors, a et b sont premiers entre eux si, et seulement si, 1 est un pgcd de a et b. Dans ce cas, tout pgcd de a et b est inversible.

En effet, si a et b sont premiers entre eux, alors tout diviseur commun est inversible. En particulier, la définition montre alors que 1 est un pgcd de a et b, puisque tout élément inversible divise 1. Tout pgcd de a et b est alors inversible par le lemme précédent. Inversement, si 1 est un pgcd de a et b, alors tout diviseur commun divise 1, et est donc inversible. Mais cela veut exactement dire que a et b sont premiers entre eux.

Revenons maintenant à l'étude de la factorisation dans un anneau. Un élément $a \in A \setminus \{0\}$ admet toujours comme diviseurs $u \in A^\times$ et va, $v \in A^\times$, puisque l'on peut écrire

$$a = u(u^{-1}a) = (va)v^{-1}.$$

Il se peut qu'il n'y en ait pas d'autres. Par exemple, les seuls diviseurs d'un nombre premier p dans $\mathbb{Z}$ sont ± 1, $\pm p$. Cela motive la définition suivante.

1.13. Définition. Un élément $\pi \in A$ est dit *irréductible* s'il est non nul, non inversible, et si pour tous $a, b \in A$, on a

$$\pi = ab \quad \Longrightarrow \quad a \in A^\times \text{ ou } b \in A^\times.$$

Cela revient à dire que π est non nul, non inversible, et que ses seuls diviseurs sont les éléments $u \in A^\times$ et $v\pi$, $v \in A^\times$.

1.14. Remarque. Si a et b sont deux éléments de A associés, alors a est irréductible dans A si, et seulement si, b l'est. Dans ce cas, on a nécessairement $b = ua$, avec $u \in A^\times$ (ce qui n'est plus le cas si a et b sont quelconques, cf. exercice 19).

En effet, supposons que a et b soient associés, i.e. $(a) = (b)$, et supposons que a soit irréductible. Comme a est non nul et non inversible, (a) est non nul et distinct de A. Il en est donc de même de (b), et b est alors non nul et non inversible. Puisque $a \in (a) = (b)$, il existe $c \in A$ tel que $a = bc$. Puisque b est non inversible et a est irréductible, on a $c \in A^\times$.

Il reste donc à montrer le fait suivant : si π est irréductible et $u \in A^\times$, alors $u\pi$ est irréductible. Cela est laissé en exercice au lecteur.

Attention ! Un anneau ne possède pas nécessairement d'éléments irréductibles, par exemple si c'est un corps.

Il y a aussi des exemples moins triviaux. Par exemple, l'anneau $\mathbb{Z}/6\mathbb{Z}$ ne possède pas d'éléments irréductibles (exercice 23). En revanche, exhiber un anneau intègre qui n'est pas un corps et ne possédant pas d'éléments irréductibles est nettement plus délicat. Un exemple d'un tel anneau est proposé dans l'exercice XXVI-14.

1.15. Exemples

(1) Un entier relatif $n \in \mathbb{Z}$ est irréductible si, et seulement si, $n = \pm p$, avec p premier.

(2) Soit $A = \mathbb{Z}[i\sqrt{3}]$. Alors, les éléments 2 et $1 \pm i\sqrt{3}$ sont irréductibles non associés deux à deux.

Remarquons tout d'abord que si $z \in A$ est inversible si, et seulement si, $|z|^2 = 1$.

En effet, si $|z|^2 = 1$, alors $z\overline{z} = 1$, et comme on a aussi $\overline{z} \in A$, on en déduit que z est inversible. Inversement, si z est inversible, alors il existe $z' \in A$ tel que $zz' = 1$, et donc $|z|^2|z'|^2 = 1$. Comme $|z|^2$ et $|z'|^2$ sont des entiers positifs, on en déduit en particulier que $|z|^2 = 1$.

Déterminons $A^\times$. Par ce qui précède, $z = a + ib\sqrt{3}$ est inversible si, et seulement si, $a^2 + 3b^2 = 1$. Nécessairement, on a $b = 0$, ce qui implique que $a = \pm 1$. Réciproquement, 1 et -1 sont inversibles. Par conséquent, on a $A^\times = \{\pm 1\}$.

Supposons maintenant que z soit égal à 2, $1 + i\sqrt{3}$ ou $1 - i\sqrt{3}$.

Alors, on a $|z|^2 = 4$, donc z est non nul et non inversible par le point précédent. De plus, si l'on a $z = z_1 z_2$, $z_i \in A$, alors $|z_1|^2 \mid 4$. On a donc

$$|z_1|^2 = 1, \ 2 \ \text{ou} \ 4,$$

selon les cas. Dans le premier cas, $z_1 \in A^\times$, et dans le dernier, $z_2 \in A^\times$ (car on a alors $|z_2|^2 = 1$). Enfin, supposons que $|z_1|^2 = 2$, et écrivons

$$z_1 = a + ib\sqrt{3}, \quad a, b \in \mathbb{Z}.$$

Alors, $a^2 + 3b^2 = 2$, et l'on a donc nécessairement $b = 0$. Mais alors, on a $a^2 = 2$, ce qui est impossible.

Ainsi, les éléments 2, $1 + i\sqrt{3}$ et $1 - i\sqrt{3}$ sont irréductibles. Il reste à voir qu'ils sont non associés deux à deux, ce qui est clair puisque les éléments inversibles de A sont ± 1.

1.16. Remarque. Si K est un corps, alors $f \in K[X]$ est irréductible si, et seulement si, il n'est pas constant et ne peut pas s'écrire comme produit de deux polynômes non constants, ce qui est la définition classique de l'irréductibilité d'un polynôme.

En effet, si f est constant, alors il est soit nul, soit inversible, car

$$K[X]^\times = K^\times = K \setminus \{0\}.$$

Dans ce cas, il n'est pas irréductible. Si $f = f_1 f_2$, avec f_1, $f_2 \in K[X]$ non constants, alors f n'est pas irréductible non plus, puisque f_1, f_2 sont alors non inversibles.

Inversement, supposons que f ne soit pas irréductible, et supposons qu'il soit non constant. Alors, en particulier il est non nul et non inversible, et donc nécessairement, on a $f = f_1 f_2$, avec f_1, $f_2 \notin K[X]^\times$ et donc non constants.

La notion de polynôme irréductible à coefficients dans un anneau quelconque est plus délicate. Nous y reviendrons dans un chapitre ultérieur.

1.17. Exemples

(1) Les polynômes irréductibles de $\mathbb{C}[X]$ sont les polynômes de degré 1.

(2) Les polynômes irréductibles de $\mathbb{R}[X]$ sont les polynômes de degré 1 et les polynômes de degré 2 de discriminant strictement négatif.

La notion d'élément irréductible nous permet de donner d'autres exemples d'éléments premiers entre eux.

1.18. Exemple. Soient π, $\pi' \in A$ deux éléments irréductibles non associés. Alors, π et π' sont premiers entre eux. En effet, soit $d \in A$ tel que $d \mid \pi$ et $d \mid \pi'$. Il existe donc des éléments b, $b' \in A$ tels que $\pi = db$ et $\pi' = db'$. Puisque π est un élément irréductible, on a $d \in A^\times$ ou $b \in A^\times$. Supposons que $b \in A^\times$. Alors, $d = b^{-1}\pi$, et $\pi' = (b^{-1}b')\pi$. Comme π est irréductible, π est non inversible. Mais, π' étant un élément irréductible, on en déduit que l'on a $u = b^{-1}b' \in A^\times$. En particulier, π et π' sont associés, ce qui contredit l'hypothèse. Ainsi, $d \in A^\times$, ce que l'on voulait vérifier.

Il n'est pas toujours aisé de démontrer qu'un élément est irréductible. Le lemme suivant donne une condition suffisante pour qu'un élément de A soit irréductible.

1.19. Lemme. *Soit A un anneau, et soit $\pi \in A$. Alors, les propriétés suivantes sont équivalentes.*

(1) *l'idéal (π) est premier non nul ;*

(2) *l'élément π est non nul, non inversible, et pour tous a, $b \in A$, on a*

$$\pi \mid ab \implies \pi \mid a \text{ ou } \pi \mid b.$$

Si de plus A est intègre, et si π vérifie une de ces deux conditions équivalentes, alors π est irréductible.

Démonstration. Soit $\pi \in A$. Supposons que (π) soit premier non nul, et soient $a, b \in A$ tels que $\pi \mid ab$. Alors, on a $ab \in (\pi)$, et puisque (π) est un idéal premier, on a $a \in (\pi)$ ou $b \in (\pi)$, c'est-à-dire $\pi \mid a$ ou $\pi \mid b$. De plus, on a $\pi \neq 0$, car $(\pi) \neq (0)$, et $\pi \notin A^\times$, car $(\pi) \neq A$.

Inversement, si π vérifie la condition (2), montrons que (π) est premier. Puisque π est non inversible et non nul, on a $(\pi) \neq (0)$, et $(\pi) \neq A$. Si maintenant $a, b \in A$ vérifient $ab \in (\pi)$, alors $\pi \mid ab$, et donc par hypothèse $\pi \mid a$ ou $\pi \mid b$, soit encore $a \in (\pi)$ ou $b \in (\pi)$. Ainsi, (π) est premier.

Supposons maintenant que A soit intègre, et que π vérifie (2) par exemple. Montrons que π est irréductible. Par hypothèse, π est non nul et non inversible. Soient $a, b \in A$ tels que $\pi = ab$. En particulier, $\pi \mid ab$ et donc $\pi \mid a$ ou $\pi \mid b$.

Supposons par exemple que $\pi \mid a$. Alors, il existe $u \in A$ tel que $a = u\pi$, et donc $\pi = u\pi b$. Comme A est intègre et $\pi \neq 0$, on obtient $1 = ub$. Par conséquent, $b \in A^\times$. Ainsi, π est irréductible. □

Les éléments vérifiant une des conditions équivalentes du lemme 1.19 méritent un nom particulier.

1.20. Définition. Soit A un anneau. Un élément $\pi \in A$ est dit *premier* s'il est non nul, non inversible, et si pour tous $a, b \in A$, on a

$$\pi \mid ab \implies \pi \mid a \text{ ou } \pi \mid b.$$

De manière équivalente, $\pi \in A$ est premier s'il engendre un idéal premier non nul de A.

1.21. Exemple. Tout nombre premier p est un élément premier de $\mathbb{Z}$.

1.22. Remarques.

(1) Si a et b sont deux éléments de A associés, alors a est premier si, et seulement si, b l'est (puisque $(a) = (b)$!).
En particulier, si π est premier et $u \in A^\times$, alors $u\pi$ est premier.

(2) Si l'anneau A est intègre, tout élément premier est irréductible, d'après le lemme 1.19. Ce n'est plus vrai si A n'est pas intègre (cf. exercice 23). De plus, un élément irréductible n'est pas nécessairement premier, comme le montre l'exemple suivant.

1.23. Exemple. Soit $A = \mathbb{Z}[i\sqrt{3}]$. Alors, 2 est un élément irréductible d'après l'exemple 1.15 (2). En revanche, il est non premier.

En effet, on a $2 \mid 4 = (1 + i\sqrt{3})(1 - i\sqrt{3})$, mais $2 \nmid (1 \pm i\sqrt{3})$.

Continuons par un lemme utile.

1.24. Lemme. *Soit $f : A \xrightarrow{\sim} B$ un isomorphisme d'anneaux. Alors, pour tout $\pi \in A$, π est irréductible si, et seulement si, $f(\pi)$ est irréductible.*

Démonstration. Supposons π irréductible. Alors, π étant non nul et f étant injective, $f(\pi)$ est non nul. De plus, f induit un isomorphisme de groupes

$$f : A^\times \xrightarrow{\sim} B^\times.$$

En particulier, puisque π n'est pas inversible, $f(\pi)$ n'est pas inversible non plus. Supposons maintenant que l'on ait

$$f(\pi) = bc, \quad b, c \in B.$$

Alors, on a $\pi = f^{-1}(b)f^{-1}(c)$. Par hypothèse, $f^{-1}(b)$ ou $f^{-1}(c)$ est inversible dans A, et ainsi b ou c est inversible dans B. Par conséquent, $f(\pi)$ est bien irréductible. Si maintenant $f(\pi)$ est irréductible, alors par le point précédent appliqué à f^{-1}, on obtient que π est irréductible, ce qui achève la démonstration. □

On peut maintenant se poser la question suivante : tout élément non nul d'un anneau A peut-il se décomposer en produit d'un élément inversible et d'éléments irréductibles ? Si oui, la décomposition est-elle unique à permutation et association près des facteurs ?

En ce qui concerne l'existence d'une telle décomposition, la réponse est non en général. On peut exhiber des exemples d'anneaux pour lesquels certains éléments n'ont pas de décomposition (cf. exercices 28 et 29), ou pis encore, des exemples d'anneaux qui ne sont pas des corps et qui ne possèdent aucun élément irréductible (cf. exercices 23 et 14). L'exemple suivant montre que l'unicité d'une décomposition n'est pas non plus assurée.

1.25. Exemple. Soit $A = \mathbb{Z}[i\sqrt{3}]$. D'après l'exemple 1.15 (2), les éléments 2, $1 + i\sqrt{3}$ et $1 - i\sqrt{3}$ sont irréductibles non associés deux à deux. Inopinément, on a

$$4 = 2^2 = (1 + i\sqrt{3})(1 - i\sqrt{3}).$$

Ainsi, l'élément 4 possède deux décompositions distinctes.

L'exemple précédent montre que dans un anneau quelconque, un élément peut avoir deux décompositions en produit d'un élément inversible et d'éléments irréductibles, sans que les éléments irréductibles mis en jeu soient associés deux à deux après une permutation convenable des facteurs. Le lemme suivant montre que les choses se passent beaucoup mieux si les éléments intervenant dans la décomposition sont premiers.

1.26. Lemme. *Soit A un anneau intègre. On suppose que l'on a une égalité*

$$u\pi_1 \cdots \pi_r = u'\pi'_1 \cdots \pi'_s,$$

où $r, s \geqslant 0$, $u, u' \in A^\times$, $\pi_1, \ldots, \pi_r$ sont premiers et $\pi'_1, \ldots, \pi'_s$ sont irréductibles. Alors, $r = s$ et il existe une permutation $\sigma \in \mathfrak{S}_r$ telle que les éléments π'_i et $\pi_{\sigma(i)}$ soient associés pour tout $i \in [\![1, r]\!]$.

Démonstration. Gardons les notations de l'énoncé. Nous voulons démontrer que $r = s$ et que chaque π_i est associé à un π'_j. Nous allons le faire par récurrence sur r. Plus précisément, l'hypothèse de récurrence sera la suivante :

(H_r) Si l'on a une égalité

$$u\pi_1 \cdots \pi_r = u'\pi'_1 \cdots \pi'_s,$$

avec $s \geqslant 0$, $u, u' \in A^\times$, $\pi_1, \ldots, \pi_r$ premiers et $\pi_1, \ldots, \pi'_s$ irréductibles, alors on a $r = s$ et chaque π_i est associé à un π'_j.

Traitons d'abord le cas $r = 0$. On a donc $u = u'\pi'_1 \cdots \pi'_s$, soit

$$u^{-1}u\pi'_1 \cdots \pi'_s = 1.$$

Supposons que $s \geqslant 1$. Alors, l'égalité précédente implique que π'_1 est inversible, ce qui est absurde puisque π'_1 est irréductible. Ainsi, $s = 0$ et (H_0) est vraie.

Supposons maintenant que (H_r) soit vraie pour un entier $r \geqslant 0$, et montrons que (H_{r+1}) est vraie. Considérons donc une égalité du type

$$u\pi_1 \cdots \pi_{r+1} = u'\pi'_1 \cdots \pi'_s,$$

avec $s \geqslant 0$.

Puisque π_1 est premier et que $\pi_1 \mid u'\pi'_1 \cdots \pi'_s$, on a $\pi_1 \mid u'$ ou π_1 divise un des π'_j. Le premier cas ne peut se produire, puisque π_1 est premier, donc non inversible. Ainsi, quitte à changer la numérotation, on peut supposer que $\pi_1 \mid \pi'_1$. On a donc

$$\pi'_1 = v\pi_1, \quad v \in A.$$

Puisque π'_1 est irréductible, on a $\pi_1 \in A^\times$ ou $v \in A^\times$. Mais, π_1 étant premier, il est non inversible et donc $v \in A^\times$. Ainsi, π_1 et π'_1 sont associés. On a donc $\pi'_1 = v\pi_1$, $v \in A^\times$, et en simplifiant par π_1 par intégrité, on obtient

$$u\pi_2 \cdots \pi_{r+1} = u'v\pi'_2\pi'_3 \cdots \pi'_s.$$

Par hypothèse de récurrence, on obtient que $s - 1 = r$ et que chaque élément irréductible π_i, $i \geqslant 2$ est associé à un π'_j, $j \geqslant 2$. Comme π_1 est associé à π'_1, cela achève la récurrence, ainsi que la démonstration. □

1.27. Remarque. Le lecteur remarquera que cette démonstration est calquée sur la démonstration de l'unicité d'une décomposition d'un entier non nul en produit de facteurs premiers et de ± 1. De plus, la conclusion du lemme montre qu'en fait, les éléments $\pi'_1, \ldots, \pi'_s$ sont premiers, puisqu'associés à des éléments premiers (cf. remarque 1.22 (1)).

Tout ce qui précède montre que les propriétés arithmétiques de $\mathbb{Z}$ ne sont plus vraies en général dans un anneau quelconque, même intègre. Nous allons maintenant montrer que dans les anneaux principaux, tout se passe pour le mieux.

2. Arithmétique des anneaux principaux

Nous commençons ce paragraphe en énonçant un résultat d'existence d'un pgcd et d'un ppcm, dont la démonstration est identique à celle de la proposition VI-3.10.

2.1. Proposition. *Soit A un anneau principal. Soient a, b, d, $m \in A$ tels que*

$$(a) + (b) = (d) \quad \text{et} \quad (a) \cap (b) = (m).$$

Alors, d est un pgcd de a et b, et m est un ppcm de a et b.

Le résultat suivant est alors une conséquence immédiate de la proposition précédente et de la remarque 1.12, et généralise le théorème de Bézout.

2.2. Corollaire. *Soit A un anneau principal.*
Deux éléments a, $b \in A$ sont premiers entre eux si, et seulement si, ils sont étrangers. Autrement dit, a et b sont premiers entre eux si, et seulement si, il existe u, $v \in A$ tels que $ua + vb = 1$.

On peut alors démontrer une généralisation du lemme de Gauss à tout anneau principal.

2.3. Lemme (Gauss). *Soit A un anneau principal. Soient a, $b \in A$ deux éléments premiers entre eux. Alors, pour tout $c \in A$, $a \mid bc \implies a \mid c$.*

Démonstration. Identique à celle du lemme II-3.3. □

Le corollaire 2.2 et le théorème XVI-3.13 permettent aussi d'obtenir le résultat suivant.

2.4. Théorème. *Soit A un anneau principal, et soient $a_1, \ldots, a_n \in A$ deux à deux premiers entre eux. Alors, le morphisme*

$$\begin{aligned} f\colon A &\longrightarrow A/(a_1) \times \cdots \times A/(a_n) \\ x &\longmapsto ([x]_1, \ldots, [x]_n) \end{aligned}$$

est surjectif, de noyau $(a_1 \cdots a_n)$. En particulier, on a un isomorphisme

d'anneaux

$$A/(a_1 \cdots a_n) \simeq A/(a_1) \times \cdots \times A/(a_n).$$

Nous allons maintenant nous intéresser de plus près aux éléments irréductibles. Le lemme suivant est fondamental, et a pas mal de conséquences intéressantes.

2.5. Lemme. *Soit A un anneau (commutatif), et soit π un élément irréductible. Alors, (π) est maximal parmi les idéaux propres principaux de A. Autrement dit, pour tout $a \in A \setminus A^\times$, on a*

$$(\pi) \subset (a) \quad \Longrightarrow \quad (\pi) = (a).$$

En particulier, dans un anneau principal, tout élément irréductible engendre un idéal maximal, et tout élément irréductible est premier.

Démonstration. Soit $\pi \in A$ un élément irréductible et soit $a \in A \setminus A^\times$, tel que $(\pi) \subset (a)$. Il existe donc $b \in A$ tel que $\pi = ab$. Puisque π est irréductible, et que $a \in A \setminus A^\times$, on a $b \in A^\times$. Mais alors, $(\pi) = (a)$, ce qu'il fallait démontrer.

Supposons maintenant que A soit principal, et soit π un élément irréductible. Comme π est en particulier non inversible, on a $(\pi) \neq A$. En particulier, il existe un idéal maximal $\mathfrak{m}$ contenant (π). Soit $a \in A$ un générateur de $\mathfrak{m}$. On a donc $(\pi) \subset (a)$, d'où $(\pi) = (a)$ par le point précédent, soit encore $(\pi) = \mathfrak{m}$. Ainsi, (π) est maximal.

En particulier, (π) est un idéal premier non nul, car π est non nul. Par définition, π est donc un élément premier. □

Notons que le fait qu'un élément irréductible d'un anneau principal soit premier n'est rien d'autre qu'une généralisation du lemme d'Euclide, bien connu dans $\mathbb{Z}$. Ce résultat est suffisamment important que nous le reformulions à part dans un lemme.

2.6. Lemme (Euclide). *Soit A un anneau principal, et soit $\pi \in A$ un élément irréductible. Alors, pour tous $a, b \in A$, on a*

$$\pi \mid ab \quad \Longrightarrow \quad \pi \mid a \ \ ou \ \ \pi \mid b.$$

Continuons en énonçant une autre conséquence du lemme 2.5.

2.7. Proposition. *Soit A un anneau principal. Alors, tout idéal premier non nul est maximal.*

Démonstration. Soit $\mathfrak{p}$ un idéal premier non nul, et soit $\pi \in A$ un générateur de $\mathfrak{p}$. Alors, π est en particulier irréductible par le lemme 1.19. D'après le lemme 2.5, $\mathfrak{p} = (\pi)$ est maximal. □

Le corollaire suivant est extrêmement utile.

2.8. Corollaire. *Soit A un anneau principal, et soit $\pi \in A \setminus \{0\}$. Les propriétés suivantes sont équivalentes :*

(1) *l'anneau $A/(\pi)$ est un corps ;*

(2) *l'anneau $A/(\pi)$ est intègre ;*

(3) *l'élément π est irréductible.*

Démonstration. D'après le théorème XVI-2.7, cela revient à démontrer les équivalences suivantes :

$$(\pi) \text{ est maximal} \iff (\pi) \text{ est premier} \iff \pi \text{ est irréductible.}$$

Comme π est non nul, tout cela est conséquence de la proposition 2.7 et du lemme 2.5. □

Nous nous intéressons ci-après à l'existence et l'unicité d'une décomposition d'un élément non nul en produit d'un inversible et d'éléments irréductibles. La première question qui vient naturellement à l'esprit est la suivante : un élément non nul et non inversible admet-il un diviseur irréductible ?

2.9. Lemme. *Soit A un anneau principal. Alors, tout élément non nul et non inversible possède un diviseur irréductible. En particulier, tout anneau principal qui n'est pas un corps possède au moins un élément irréductible.*

Démonstration. On suppose que A n'est pas un corps, sinon le résultat est trivial. En particulier, A possède au moins un idéal maximal. Montrons maintenant que tout générateur d'un idéal maximal est irréductible. Soit $\mathfrak{m}$ un idéal maximal. En particulier, $\mathfrak{m}$ est un idéal premier, non nul (sinon (0) serait maximal et A serait un corps). Écrivons $\mathfrak{m} = (\pi)$. Alors, π est premier, donc irréductible puisqu'un anneau principal est intègre. Cela démontre déjà le second point.

Soit maintenant $a \in A$ un élément non nul et non inversible. Puisque a est non inversible, on a $(a) \neq A$. L'idéal (a) est donc contenu dans un idéal maximal $\mathfrak{m}$ par le théorème de Krull. D'après ce qui précède, on a $\mathfrak{m} = (\pi)$, où π est irréductible. Mais alors, comme $(a) \subset (\pi)$, on a $\pi \mid a$, et π est un diviseur irréductible de A. Cela achève la démonstration. □

On peut alors énoncer et démontrer le théorème suivant.

2.10. Théorème. *Soit A un anneau principal.*

Alors, tout élément non nul $a \in A$ peut s'écrire

$$a = u\pi_1 \cdots \pi_r,$$

où $\pi_1, \ldots, \pi_r$ sont irréductibles, et $u \in A^\times$. De plus, cette écriture est unique à permutation et association des facteurs près.

Autrement dit, si l'on a deux décompositions

$$a = u\pi_1 \cdots \pi_r = u'\pi'_1 \cdots \pi'_s,$$

alors $r = s$ et il existe $\sigma \in \mathfrak{S}_r$ telle que π_i et $\pi'_{\sigma(i)}$ soient associés pour tout entier $i \in [\![1, r]\!]$.

Démonstration. Soit $a \in A \setminus \{0\}$. Commençons par démontrer l'existence d'une décomposition de la forme voulue. Si $a \in A^\times$, c'est clair. On suppose donc que A n'est pas un corps, et que a est non nul et non inversible. D'après le lemme précédent, a possède au moins un diviseur irréductible $\pi_1 \in A$. Écrivons $a = \pi_1 a_1$, $a_1 \in A$. Si a_1 est inversible, on s'arrête. Si a_1 n'est pas inversible, on a $(a) \subsetneq (a_1)$. En effet, puisque $a_1 \mid a$, on a $(a) \subset (a_1)$. De plus, si $(a) = (a_1)$, on a $a = ua_1$, pour un certain $u \in A^\times$ (car A est intègre). Mais alors $a = \pi_1 a_1 = ua_1$. Remarquons maintenant que $a_1 \neq 0$, car $a \neq 0$. Par intégrité, on a $\pi_1 = u \in A^\times$, d'où une contradiction.

Puisque a_1 est non nul et non inversible, il possède alors un diviseur irréductible π_2. On écrit alors $a_1 = \pi_2 a_2$. Si a_2 est inversible, on s'arrête. Sinon on a comme précédemment $(a_1) \subsetneq (a_2)$, et l'on continue. Supposons que l'élément a_n construit à l'étape n ne soit jamais inversible. Par récurrence, on construit alors une chaîne infinie d'idéaux

$$(a) \subsetneq (a_1) \subsetneq \cdots \subsetneq (a_n) \subsetneq \ldots,$$

où $a_n \in A \setminus A^\times$, pour tout $n \geqslant 1$.

Posons $a_0 = a$ et soit $\mathfrak{a} = \bigcup_{n \geqslant 0} (a_n)$. Alors, $\mathfrak{a}$ est un idéal de A. Puisque A est un anneau principal, il existe un élément $x \in A$ tel que $\mathfrak{a} = (x)$. En particulier, $x \in (a_j)$ pour un certain entier $j \geqslant 0$, et donc

$$(x) \subset (a_j) \subset \mathfrak{a} = (x),$$

d'où $(x) = (a_j)$. Mais alors, on obtient

$$(x) = (a_j) \subsetneq (a_{j+1}) \subset (x),$$

et donc $(x) \subsetneq (x)$, d'où une contradiction.

Ainsi, il existe $n \geqslant 1$ tel que $a_{n+1} \in A^\times$. En posant $u = a_{n+1}$, on obtient alors $a = u\pi_1 \ldots \pi_n$, ce qui est la décomposition cherchée. L'unicité découle

du lemme 1.26, puisque tout élément irréductible est premier d'après le lemme 2.5. □

Il peut être intéressant d'avoir une réelle unicité de la décomposition d'un élément en produits d'irréductibles. Pour cela, on introduit la notion de système complet de représentants irréductibles.

2.11. Définition. Soit A un anneau. Un *système complet de représentants irréductibles* (s.c.r.i.) est un sous-ensemble $\mathscr{P}$ de A tel que

(1) tout élément de $\mathscr{P}$ est irréductible ;

(2) tout élément irréductible de A est associé à un unique élément de $\mathscr{P}$;

2.12. Exemples

(1) $A = \mathbb{Z}$, $\mathscr{P} = \{2, 3, 5, 7, 11, \ldots\}$.

(2) $A = \mathbb{Z}$, $\mathscr{P} = \{-2, -3, -5, -7, \ldots\}$.

(3) $A = \mathbb{C}[X]$, $\mathscr{P} = \{X - a \mid a \in \mathbb{C}\}$.

(4) $A = K[X]$ (où K est un corps), $\mathscr{P}$ est l'ensemble des polynômes irréductibles unitaires.

Un tel s.c.r.i. existe toujours. En effet, il suffit de munir l'ensemble des éléments irréductibles de A de la relation d'équivalence « être associés », et de prendre un représentant dans chaque classe d'équivalence.

Une fois choisi un s.c.r.i. $\mathscr{P}$, lorsque A est principal, on peut toujours écrire chaque diviseur irréductible intervenant dans une décomposition d'un élément $a \in A$ non nul comme le produit d'un élément inversible et d'un élément de $\mathscr{P}$. En regroupant les termes, on voit que a s'écrit comme produit d'un élément inversible et de puissances d'éléments de $\mathscr{P}$.

Le théorème 2.10 peut alors se reformuler comme suit.

2.13. Théorème. *Soit A un anneau principal, et soit $\mathscr{P}$ un s.c.r.i. de A. Alors, tout élément $a \in A \setminus \{0\}$ s'écrit sous la forme*

$$a = u \prod_{\pi \in \mathscr{P}} \pi^{n_\pi},$$

où $u \in A^\times$ et les entiers $n_\pi \in \mathbb{N}$ sont tous nuls, sauf un nombre fini. De plus, l'élément $u \in A$ et la famille d'entiers $(n_\pi)_{\pi \in \mathscr{P}}$ sont uniques.

La démonstration de l'équivalence de ces deux énoncés est laissée à la sagacité du lecteur.

Grâce à ce dernier théorème, on obtient alors une conséquence du théorème chinois, qui généralise le corollaire XVII-1.2.

2.14. Théorème. *Soit A un anneau principal, soit $\mathscr{P}$ un s.c.r.i. de A et soit $a \in A \setminus \{0\}$.*

Si $a = u \prod_{\pi \in \mathscr{P}} \pi^{n_\pi}$, on a alors un isomorphisme d'anneaux

$$A/(a) \simeq \prod_{\pi \in \mathscr{P}} A/(\pi^{n_\pi}).$$

Démonstration. Il suffit de remarquer les éléments π^{n_π} sont premiers entre eux deux à deux en utilisant le lemme d'Euclide par exemple, et d'appliquer le théorème 2.4. □

Pour finir ce paragraphe, notons que le théorème 2.13 donne en théorie un moyen de calculer un pgcd de deux éléments a, b non nuls d'un anneau principal, une fois choisi un s.c.r.i. $\mathscr{P}$. Il suffit en effet d'écrire la décomposition de a et b en produit d'un inversible et de puissances d'éléments de $\mathscr{P}$, et de procéder comme dans $\mathbb{Z}$ (cf. corollaire II-4.7).

Néanmoins, cela suppose que l'on sache expliciter décomposition en produits d'irréductibles d'un élément donné, ce qui suppose aussi implicitement d'avoir un test fiable d'irréductibilité. Cela est déjà impossible de manière efficace dans $\mathbb{Z}$. L'autre possibilité pour calculer un pgcd et un ppcm de a et b serait de savoir trouver explicitement un générateur des idéaux (a, b) et $(a) \cap (b)$, ce qui est un problème tout aussi difficile.

En revanche, dans $\mathbb{Z}$, on peut calculer un pgcd de manière systématique en utilisant une successions de divisions euclidiennes, via l'algorithme d'Euclide.

Nous allons maintenant nous intéresser aux anneaux possédant une division euclidienne, et voir que tout cela se généralise.

3. Anneaux euclidiens

3.1. Définition. Un anneau *euclidien* est un couple (A, δ), où A est un anneau intègre et $\delta : A \setminus \{0\} \longrightarrow \mathbb{N}$ est une application telle que, pour tout $a \in A$, et pour tout $b \in A \setminus \{0\}$, il existe q, $r \in A$ tels que

$$a = qb + r, \quad r = 0 \text{ ou } \delta(r) < \delta(b).$$

Attention ! On ne demande pas l'unicité de q et r.

Lorsque δ est spécifiée, on dit simplement que A est euclidien.

3.2. Exemples

(1) L'anneau $\mathbb{Z}$ est euclidien (pour la valeur absolue).
(2) Si K est un corps, K est euclidien pour la fonction constante 1.
(3) Si K est un corps, $K[X]$ est euclidien pour le degré des polynômes, d'après le théorème XV-6.10 et la remarque qui suit.

L'intérêt des anneaux euclidiens est donné a priori par le résultat suivant.

3.3. Proposition. *Tout anneau euclidien est principal.*

Démonstration. Supposons (A, δ) euclidien. Alors, A est intègre. Soit $\mathfrak{a}$ un idéal de A. Si $\mathfrak{a} = (0)$, il n'y a rien à faire. On peut donc supposer $\mathfrak{a}$ non nul. L'ensemble $\{\delta(a) \mid a \in \mathfrak{a} \setminus \{0\}\}$ est alors une partie non vide de $\mathbb{N}$, et admet donc un plus petit élément. Soit $a_0 \in \mathfrak{a} \setminus \{0\}$ réalisant ce minimum. On va montrer que $\mathfrak{a} = (a_0)$. Puisque $\mathfrak{a}$ est un idéal de A contenant a_0, on a $(a_0) \subset \mathfrak{a}$. Réciproquement, soit $a \in \mathfrak{a}$. Écrivons

$$a = qa_0 + r, \quad \text{avec } r = 0 \text{ ou } \delta(r) < \delta(a_0).$$

Puisque a et a_0 sont des éléments de $\mathfrak{a}$, on a $r = a - qa_0 \in \mathfrak{a}$. Si r était non nul, on aurait alors $r \in \mathfrak{a} \setminus \{0\}$ et $\delta(r) < \delta(a_0)$, ce qui contredirait le choix de a_0. Par conséquent, $r = 0$ et l'on a $a = qa_0 \in (a_0)$. Cela montre l'inclusion manquante. □

3.4. Exemple. On retrouve le fait que $\mathbb{Z}$ et $K[X]$ (K corps) sont des anneaux principaux. Le lecteur attentif notera d'ailleurs que les démonstrations de ces deux faits sont identiques à la précédente.

On continue en donnant quelques exemples d'anneaux euclidiens.

3.5. Lemme. *Si $A = \mathbb{Z}[i]$, $\mathbb{Z}[i\sqrt{2}]$ et $\mathbb{Z}[j]$, alors A est euclidien pour l'application*

$$\begin{aligned} \delta : A \setminus \{0\} &\longrightarrow \mathbb{N} \\ z &\longmapsto |z|^2. \end{aligned}$$

De même, si $A = \mathbb{Z}[\sqrt{2}]$, alors A est euclidien pour l'application

$$\begin{aligned} \delta : A \setminus \{0\} &\longrightarrow \mathbb{N} \\ a + b\sqrt{2} &\longmapsto |a^2 - 2b^2|. \end{aligned}$$

Démonstration. Tous ces anneaux sont intègres, puisque ce sont des sous-anneaux de $\mathbb{C}$.

Supposons tout d'abord que $A = \mathbb{Z}[i]$, $\mathbb{Z}[i\sqrt{2}]$ ou $\mathbb{Z}[\sqrt{2}]$, et posons

$$\omega = i, \ i\sqrt{2} \text{ ou } \sqrt{2},$$

selon les cas. Tout élément de A s'écrit donc de manière unique sous la

forme

$$z = a + b\omega, \quad a, b \in \mathbb{Z}.$$

Posons $B = \mathbb{Q}[\omega]$. Tout élément de B s'écrit de manière unique sous la forme

$$z = u + v\omega, \quad u, v \in \mathbb{Q}.$$

On définit alors $z^* = u - v\omega \in B$. Dans tous les cas, pour tout $z \in B$, on a

$$zz^* = u^2 - dv^2 \in \mathbb{Q}, \quad \text{avec } d \in \{-1, 1, 2\},$$

et $zz^* = 0$ si, et seulement si, $z = 0$. On étend l'application δ en une application de B dans $\mathbb{Q}$, encore notée δ par abus, en posant

$$\begin{array}{rccc} \delta : & B & \longrightarrow & \mathbb{Q} \\ & a + b\omega & \longmapsto & |a^2 - db^2|, \end{array}$$

où $d = -1$, -2 ou 2.

On remarque alors que

$$\delta(z) = |zz^*|, \quad \text{pour tout } z \in B.$$

Remarquons aussi que pour tous z_1, $z_2 \in A$, on a $(z_1 z_2)^* = z_1^* z_2^*$. Cela résulte en effet d'un simple calcul laissé au lecteur. En particulier, pour tous z_1, $z_2 \in B$, on a

$$\delta(z_1 z_2) = |z_1 z_2 z_1 z_2^*| = |z_1 z_2||z_1^* z_2^*| = \delta(z_1)\delta(z_2).$$

Notons enfin que, si $z \in A \setminus \{0\}$, $\delta(z)$ est un entier strictement positif.

Supposons maintenant que $A = \mathbb{Z}[j]$ et soit $B = \mathbb{Q}[j]$. Dans ce cas, δ est simplement le carré du module, et s'étend bien entendu à B. De plus, pour tous u, $v \in \mathbb{Q}$, on a

$$\delta(u+vj) = |u+vj|^2 = \left|\left(u - \frac{v}{2}\right) + iv\frac{\sqrt{3}}{2}\right|^2 = \left(u - \frac{v}{2}\right)^2 + \frac{3v^2}{4} = u^2 - uv + v^2.$$

En particulier, si $z \in A$, $\delta(z)$ est un entier positif, qui est nul si, et seulement si, $z = 0$. Notons également que A et B sont stables par conjugaison complexe.

Pour harmoniser les notations avec les autres cas, si $z \in B$, on note $z^* \in B$ son conjugué. On note également $\omega = j$ dans ce cas.

Passons maintenant à la démonstration proprement dite. Soient z_1, $z_2 \in A$, avec $z_2 \neq 0$. Alors, on a

$$\frac{z_1}{z_2} = \frac{z_1 z_2^*}{z_2 z_2^*}.$$

Comme $z_2 z_2^* \in \mathbb{Z} \setminus \{0\}$ et que $z_1 z_2^* \in A$, on peut donc écrire $\frac{z_1}{z_2} = u + v\omega$, avec u, $v \in \mathbb{Q}$. Soient m, $n \in \mathbb{Z}$ deux entiers vérifiant

$$|u - m| \leqslant \frac{1}{2} \quad \text{et} \quad |v - n| \leqslant \frac{1}{2}.$$

Posons $q = m + n\omega$ et $r = z_1 - qz_2 = \big(\frac{z_1}{z_2} - q\big)z_2$. On a alors

$$\delta(r) = \delta\big(((u-m) + (v-n)\omega)z_2\big) = \delta\big((u-m) + (v-n)\omega\big)\delta(z_2).$$

Nous voulons maintenant montrer que $\delta\big((u-m) + (v-n)\omega\big) < 1$.

Si $A = \mathbb{Z}[i]$, $\mathbb{Z}[i\sqrt{2}]$ ou $\mathbb{Z}[\sqrt{2}]$, on a

$$\delta\big((u-m) + (v-n)\omega\big) = |(u-m)^2 - d(v-n)^2|,$$

avec $d = -1$, -2 ou 2. Dans les trois cas, on a donc

$$\delta\big((u-m) + (v-n)\omega\big) \leqslant (u-m)^2 + 2(v-n)^2 \leqslant \frac{1}{4} + \frac{1}{2} < 1.$$

Si $A = \mathbb{Z}[j]$, on a

$$\delta\big((u-m) + (v-n)\omega\big) \leqslant |u-m|^2 + |u-m|^2 + |u-m||v-n| \leqslant \frac{3}{4} < 1.$$

Ainsi, dans tous les cas, on a $z_1 = qz_2 + r$ avec, soit $r = 0$, soit $\delta(r) < \delta(z_2)$, ce qui démontre que (A, δ) est euclidien. □

3.6. Remarque. La démonstration précédente nous donne un moyen de calculer un quotient dans chaque cas. Si z_1, $z_2 \in A$, avec $z_2 \neq 0$, on écrit

$$\frac{z_1}{z_2} = u + vi, \quad u + vj, \quad u + iv\sqrt{2}, \quad u + v\sqrt{2}, \qquad u, v \in \mathbb{Q}$$

selon le cas dans lequel on se trouve. On choisit ensuite m, $n \in \mathbb{Z}$ vérifiant

$$|u-m| \leqslant \frac{1}{2}, \quad |v-n| \leqslant \frac{1}{2}.$$

Alors, un quotient q de z_1 par z_2 est respectivement donné par

$$q = m + ni, \quad m + nj, \quad m + in\sqrt{2}, \quad m + n\sqrt{2}.$$

On pose ensuite $r = z_1 - qz_2$.

Remarquons au passage que l'on pourrait avoir plusieurs choix possibles pour m ou n, dans le cas ou $u = \frac{1}{2}$ ou $v = \frac{1}{2}$.

Nous allons maintenant généraliser l'algorithme d'Euclide bien connu pour l'anneau $\mathbb{Z}$ à tout anneau euclidien.

Algorithme d'Euclide. Soit (A, δ) un anneau euclidien, soient a et b deux éléments non nuls. Tant que c'est possible (i.e. tant que l'on ne divise pas par 0), on définit des éléments q_i et r_{i+1} comme suit par divisions euclidiennes successives en posant :

$$r_{-1} = a, \quad r_0 = b, \quad r_{i-1} = q_i r_i + r_{i+1}.$$

En particulier, si $r_i \neq 0$, on a soit $r_{i+1} = 0$, soit $\delta(r_{i+1}) < \delta(r_i)$.

Remarquons que q_i n'est défini que pour $i \geqslant 0$.

Pour toute valeur $i \geqslant -1$ pour laquelle $r_i \neq 0$, on peut alors définir des éléments u_i et v_i par récurrence en posant :

$$\begin{aligned} u_{-1} &= 1, \quad u_0 = 0, \quad u_{i+1} = -q_i u_i + u_{i-1}, \\ v_{-1} &= 0, \quad v_0 = 1, \quad v_{i+1} = -q_i v_i + v_{i-1}. \end{aligned}$$

3.7. Proposition. *Il existe un entier $i \geqslant 1$ tel que $r_i = 0$. Si $n \geqslant 0$ est le plus grand entier tel que $r_n \neq 0$, on pose*

$$d = r_n, \quad u = u_n, \quad v = v_n.$$

Alors, d est un pgcd de a et b, et l'on a la relation de Bézout

$$ua + vb = d.$$

Démonstration. Supposons que r_i soit non nul pour tout $i \geqslant -1$. Par construction, la suite $\big(\delta(r_i)\big)_{i \geqslant -1}$ est alors une suite d'entiers naturels strictement décroissante. Une telle suite n'existant pas, on obtient une contradiction. Ainsi, il existe un entier [2] $i \geqslant 1$ tel que $r_i = 0$. Soit $n \geqslant 0$ le plus grand entier tel que r_n soit non nul. On a donc $r_{n+1} = 0$, et l'on dispose de suites finies d'éléments

$$(r_i)_{i \in [\![-1, n+1]\!]}, \quad (q_i)_{i \in [\![0, n+1]\!]}, \quad (u_i)_{i \in [\![-1, n+1]\!]}, \quad (v_i)_{i \in [\![-1, n+1]\!]}.$$

Remarquons que les égalités définissant $(r_i)_{i \in [\![-1, n]\!]}$ se récrivent

$$\begin{pmatrix} r_{i-1} \\ r_i \end{pmatrix} = M_i \begin{pmatrix} r_i \\ r_{i+1} \end{pmatrix},$$

avec $M_i = \begin{pmatrix} q_i & 1 \\ 1 & 0 \end{pmatrix}$, pour tout $i \in [\![0, n]\!]$.

En particulier, la matrice produit $M = M_0 \cdots M_n$ est à coefficients dans l'anneau A. Or, on a

$$\begin{pmatrix} r_{-1} \\ r_0 \end{pmatrix} = M \begin{pmatrix} r_n \\ r_{n+1} \end{pmatrix},$$

soit

$$\begin{pmatrix} a \\ b \end{pmatrix} = M \begin{pmatrix} d \\ 0 \end{pmatrix}.$$

On en déduit immédiatement que $d \mid a$ et $d \mid b$.

Remarquons que l'on a également

$$\begin{pmatrix} u_{i-1} \\ u_i \end{pmatrix} = M_i \begin{pmatrix} u_i \\ u_{i+1} \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} v_{i-1} \\ v_i \end{pmatrix} = M_i \begin{pmatrix} v_i \\ v_{i+1} \end{pmatrix},$$

pour tout $i \in [\![0, n]\!]$.

2. Rappelons que $r_{-1} = a$ et $r_0 = b$ sont supposés non nuls.

Ainsi,
$$\begin{pmatrix} u_{-1} \\ u_0 \end{pmatrix} = M \begin{pmatrix} u_n \\ u_{n+1} \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} v_{-1} \\ v_0 \end{pmatrix} = M \begin{pmatrix} v_n \\ v_{n+1} \end{pmatrix}.$$

Matriciellement, cela se traduit par
$$\mathrm{I}_2 = M \begin{pmatrix} u_n & v_n \\ u_{n+1} & v_{n+1} \end{pmatrix}.$$

En particulier, $M \in \mathrm{GL}_2(K_A)$ et l'on a $M^{-1} = \begin{pmatrix} u_n & v_n \\ u_{n+1} & v_{n+1} \end{pmatrix}$.

L'égalité $\begin{pmatrix} a \\ b \end{pmatrix} = M \begin{pmatrix} d \\ 0 \end{pmatrix}$ nous donne alors
$$\begin{pmatrix} d \\ 0 \end{pmatrix} = M^{-1} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} u_n & v_n \\ u_{n+1} & v_{n+1} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix}.$$

En particulier, on a $au_n + bv_n = d$, soit $ua + vb = d$ avec les notations du lemme. Mais alors, si $c \in A$ est un diviseur commun à a et b, il divise également $ua + vb$, soit $c \mid d$. Cela démontre que d est un pgcd de a et b, et achève la démonstration. □

Nous décrivons maintenant un moyen pratique de faire fonctionner l'algorithme.

Disposition manuelle. On remarque que l'on a en fait
$$r_{-1} = a, \quad r_0 = b, \quad r_{i+1} = r_{i-1} - q_i r_i,$$
$$u_{-1} = 1, \quad u_0 = 0, \quad u_{i+1} = u_{i-1} - q_i u_i,$$
$$v_{-1} = 0, \quad v_0 = 1, \quad v_{i+1} = v_{i-1} - q_i v_i.$$

On commence par remplir un tableau à quatre colonnes (ou seulement deux colonnes si l'on n'est pas intéressé par le calcul d'une relation de Bézout) comme suit :

$-q$	r	u	v
	a	1	0
	b	0	1

Expliquons le procédé sur un exemple. On prend $A = \mathbb{Z}$, $a = 32$, $b = 7$. On part donc du tableau suivant :

$-q$	r	u	v
	32	1	0
	7	0	1

Remarquons que q_{-1} n'est pas défini.

L'étape suivante consiste à calculer q_0. On dispose d'un algorithme pour le faire dans les anneaux

$$\mathbb{Z}, \quad K[X], \quad \mathbb{Z}[i], \quad \mathbb{Z}[j], \quad \mathbb{Z}[i\sqrt{2}], \quad \text{et} \quad \mathbb{Z}[\sqrt{2}],$$

mais cela peut être plus délicat dans le cas d'un anneau euclidien quelconque.

Dans notre exemple, $32 = 4\times7 + 4$, et donc $q_0 = 4$. On remplit alors le tableau comme suit :

$-q$	r	u	v
	32	1	0
-4	7	0	1

Pour calculer r_1, u_1, v_1, on multiplie les éléments des colonnes 2, 3 et 4 de la dernière ligne par l'élément de la colonne 1, et l'on ajoute aux éléments correspondants dans la ligne du dessus.

Dans notre exemple, $r_1 = -4\times7 + 32$, $u_1 = -4\times0 + 1$ et $v_1 = -4\times1 + 0$. On obtient

$-q$	r	u	v
	32	1	0
-4	7	0	1
	4	1	-4

Pour calculer les valeurs suivantes, on considère les deux dernières lignes, et l'on calcule le nouveau quotient en effectuant la division des deux dernières valeurs de la colonne des restes successifs (i.e. la colonne 2).

Dans notre exemple, on doit diviser 7 par 4. On a $7 = 1 \times 4 + 3$, et le nouveau quotient est 1. On obtient

$-q$	r	u	v
	32	1	0
-4	7	0	1
-1	4	1	-4

On répète le procédé pour calculer les valeurs de la ligne suivante. La nouvelle valeur de r_i est $-1\times4 + 7$, celle de u_i est $-1\times1 + 0$, et celle de v_i est $-1\times(-4) + 1$.

On obtient donc

$-q$	r	u	v
	32	1	0
-4	7	0	1
-1	4	1	-4
	3	-1	5

Pour trouver le nouveau quotient, on divise 4 par 3 : on a $4 = 1\times3 + 1$, et ainsi le nouveau quotient est 1. On a ainsi

$-q$	r	u	v
	32	1	0
-4	7	0	1
-1	4	1	-4
-1	3	-1	5

La nouvelle valeur de r_i est $-1\times3+4$, celle de u_i est $-1\times(-1)+1$, et celle de v_i est $-1\times5+(-4)$. On a donc

$-q$	r	u	v
	32	1	0
-4	7	0	1
-1	4	1	-4
-1	3	-1	5
	1	2	-9

On constate alors que la valeur du reste suivant est 0, puisque l'on doit diviser 3 par 1. On arrête donc ici l'algorithme. Ainsi, un pgcd de 32 et 7 est 1 (la dernière valeur de r_i dans le tableau) et on a $u = 2$ et $v = -9$.

On a bien $32\times2 + 7\times(-9) = 64 - 63 = 1$.

Donnons maintenant un exemple dans $\mathbb{Z}[i]$.

3.8. Exemple. Soit $A = \mathbb{Z}[i]$, et soient $a = 12 + 45i$ et $b = 4 + 12i$. On part donc du tableau suivant :

$-q$	r	u	v
	$12+45i$	1	0
	$4+12i$	0	1

On a ensuite
$$\frac{12+45i}{4+12i} = \frac{147}{40} + \frac{9}{40}i.$$

Un quotient est donc donné par $q_0 = 4$, et l'on obtient donc

$-q$	r	u	v
	$12+45i$	1	0
-4	$4+12i$	0	1
	$-4-3i$	1	-4

On a maintenant
$$\frac{4+12i}{-4-3i} = -\frac{52}{25} - \frac{36}{25}i,$$

d'où $q_1 = -2 - i$.

Tous calculs faits, le tableau se complète comme suit :

$-q$	r	u	v
	$12+45i$	1	0
-4	$4+12i$	0	1
$2+i$	$-4-3i$	1	-4
	$-1+2i$	$2+i$	$-7-4i$

On a ensuite

$$\frac{-4-3i}{-1+2i} = -\frac{2}{5} + \frac{11}{5}i,$$

soit $q_2 = 2i$. On obtient donc

$-q$	r	u	v
	$12+45i$	1	0
-4	$4+12i$	0	1
$2+i$	$-4-3i$	1	-4
$-2i$	$-1+2i$	$2+i$	$-7-4i$
	$-i$	$3-4i$	$-12+14i$

Comme $-i$ est inversible, le prochain reste sera nécessairement nul. Un pgcd de a et b est donc $-i$, et l'on a bien la relation

$$(3-4i)(12+45i) + (-12+14i)(4+12i) = -i.$$

Nous allons maintenant revenir au problème d'existence et d'unicité d'une décomposition en irréductibles dans un cadre plus général. On commence par étudier brièvement une classe d'anneaux dans lesquels l'existence d'une telle décomposition est assurée.

4. Anneaux noethériens

Commençons par une définition.

4.1. Définition. Un anneau commutatif A est dit *noethérien* si tout idéal de A peut être engendré par un nombre fini d'éléments.

4.2. Exemples

(1) Un anneau principal est noethérien.

(2) Soit $d \in \mathbb{Z}$ un entier non nul sans facteurs carrés. Alors, $\mathbb{Z}[\sqrt{d}]$ est noethérien.

En effet, soit $\mathfrak{a}$ un idéal de A. Pour tout $a \in \mathfrak{a}$, on peut écrire

$$a = m_a + n_a\sqrt{d}, \quad m_a,\, n_a \in \mathbb{Z}.$$

Soit $\mathfrak{b}$ l'idéal de $\mathbb{Z}$ engendré par les entiers m_a, $a \in \mathfrak{a}$. Comme $\mathbb{Z}$ est principal, $\mathfrak{b}$ est engendré par un entier $m \in \mathbb{Z}$. On a ainsi

$$a = mm'_a + n_a\sqrt{d}, \quad \text{avec } m'_a \in \mathbb{Z}.$$

Puisque $m \in \mathfrak{b}$, il existe $a_1, \ldots, a_r \in \mathfrak{a}$ et $u_1, \ldots, u_r \in \mathbb{Z}$ tels que

$$m = u_1 m_{a_1} + \cdots + u_r m_{a_r}.$$

Posons $y_1 = u_1 a_1 + \cdots + u_r a_r \in \mathfrak{a}$. Alors, on a $y_1 = m + n_{y_1}\sqrt{d}$, et ainsi pour tout $a \in \mathfrak{a}$, on a

$$a - m'_a y_1 = n'_a\sqrt{d}, \quad n'_a \in \mathbb{Z}.$$

Remarquons que, puisque $y_1 \in \mathfrak{a}$, on a $n'_a\sqrt{d} \in \mathfrak{a}$ pour tout $a \in \mathfrak{a}$. Soit $\mathfrak{c}$ l'idéal de $\mathbb{Z}$ engendré par les entiers n'_a, et soit $n' \in \mathbb{Z}$ un générateur de $\mathfrak{c}$. Alors, il existe $b_1, \ldots, b_s \in \mathfrak{a}$ et $v_1, \ldots, v_s \in \mathbb{Z}$ tels que

$$n' = v_1 n'_{b_1} + \cdots + v_s n'_{b_s}.$$

En multipliant par $\sqrt{d}$, on obtient que $y_2 = n'\sqrt{d} \in \mathfrak{a}$. Or, on a

$$n'_a = n' n''_a, \quad n''_a \in \mathbb{Z},$$

par définition. On peut donc écrire

$$a = m'_a y_1 + n''_a y_2,$$

pour tout $a \in \mathfrak{a}$. Ainsi, $\mathfrak{a}$ est engendré par les éléments y_1, y_2 (qui sont bien des éléments de $\mathfrak{a}$), et $\mathfrak{a}$ est donc de type fini.

(3) Plus généralement, le même genre d'arguments montre qu'un anneau qui est de type fini comme groupe abélien est noethérien.

On continue par quelques résultats généraux sur les anneaux noethériens.

4.3. Lemme. *Soit $f : A \longrightarrow B$ un morphisme d'anneaux surjectif. Si A est noethérien, alors B est noethérien.*

En particulier :

(1) *tout quotient d'un anneau noethérien est noethérien ;*

(2) *si A et B sont deux anneaux isomorphes, alors A est noethérien si, et seulement si, B est noethérien.*

Démonstration. Supposons A noethérien. Soit $\mathfrak{b}$ un idéal de B. Comme f est surjective, on a $\mathfrak{b} = f\big(f^{-1}(\mathfrak{b})\big)$. Or, $f^{-1}(\mathfrak{b})$ est un idéal de A, donc

engendré par un nombre fini d'éléments $a_1, \ldots, a_n$. Mais alors, f étant surjective, $\mathfrak{b}$ est engendré par $f(a_1), \ldots, f(a_n)$. Ainsi, $\mathfrak{b}$ est de type fini, et l'anneau B est noethérien.

Si maintenant $\mathfrak{a}$ est un idéal de A, il suffit d'appliquer le point précédent à la projection canonique $A \longrightarrow A/\mathfrak{a}$ pour obtenir la noethérianité de $A/\mathfrak{a}$.

Enfin, si $f : A \xrightarrow{\sim} B$ est un isomorphisme, il suffit d'appliquer le premier point à f et f^{-1} pour obtenir (2). Cela achève la démonstration. □

4.4. Lemme. *Soient A_1 et A_2 deux anneaux. Alors, tout idéal de $A_1 \times A_2$ est de la forme $\mathfrak{a}_1 \times \mathfrak{a}_2$, où $\mathfrak{a}_i$ est un idéal de A_i.*

En particulier, si tout idéal de A_i est de type fini, resp. principal ($i = 1, 2$), alors tout idéal de $A_1 \times A_2$ est de type fini, resp. principal.

Démonstration. Soit $\mathfrak{a}$ un idéal de $A_1 \times A_2$. Posons

$$\mathfrak{a}_1 = \{x_1 \in A_1 \mid (x_1, 0) \in \mathfrak{a}\},$$

ainsi que

$$\mathfrak{a}_2 = \{x_2 \in A_2 \mid (0, x_2) \in \mathfrak{a}\}.$$

On vérifie aisément $\mathfrak{a}_i$ est un idéal de A_i pour $i = 1, 2$. Montrons que l'on a

$$\mathfrak{a} = \mathfrak{a}_1 \times \mathfrak{a}_2.$$

Si $(x_1, x_2) \in \mathfrak{a}_1 \times \mathfrak{a}_2$, on a donc $(x_1, x_2) = (x_1, 0) + (0, x_2) \in \mathfrak{a}$, puisque $\mathfrak{a}$ est stable par addition. Ainsi, $\mathfrak{a}_1 \times \mathfrak{a}_2 \subset \mathfrak{a}$. Si maintenant $(x_1, x_2) \in \mathfrak{a}$, on a aussi $(x_1, x_2)(1, 0) \in \mathfrak{a}$ et $(x_1, x_2)(0, 1) \in \mathfrak{a}$, c'est-à-dire $(x_1, 0), (0, x_2) \in \mathfrak{a}$. On a donc $x_1 \in \mathfrak{a}_1$ et $x_2 \in \mathfrak{a}_2$. Par conséquent, $(x_1, x_2) \in \mathfrak{a}_1 \times \mathfrak{a}_2$. On a ainsi $\mathfrak{a} \subset \mathfrak{a}_1 \times \mathfrak{a}_2$, et cela montre que $\mathfrak{a} = \mathfrak{a}_1 \times \mathfrak{a}_2$.

Montrons la dernière partie. Si $a_1, \ldots, a_r$ engendrent $\mathfrak{a}_1$ et $b_1, \ldots, b_s$ engendrent $\mathfrak{a}_2$, il est clair que $(a_1, 0), \ldots, (a_r, 0), (0, b_1), \ldots, (0, b_s)$ engendrent l'idéal $\mathfrak{a}_1 \times \mathfrak{a}_2$.

Enfin, si $\mathfrak{a}_i = a_i A_i$, $i = 1, 2$, alors $\mathfrak{a}_1 \times \mathfrak{a}_2 = (a_1, a_2)(A_1 \times A_2)$. Cela achève la démonstration. □

4.5. Corollaire. *Soient $A_1, \ldots, A_n$ des anneaux. Alors, $A_1 \times \cdots \times A_n$ est noethérien si, et seulement si, $A_1, \ldots, A_n$ sont noethériens.*

Démonstration. Si $A_1, \ldots, A_n$ sont noethériens, le lemme 4.4 et une récurrence immédiate montrent alors que $A_1 \times \cdots \times A_n$ est noethérien. Inversement, si $A_1 \times \cdots \times A_n$ est noethérien, le lemme 4.3, appliqué aux diverses projections, montre que $A_1, \ldots, A_n$ sont noethériens. □

On va maintenant donner des caractérisations variées des anneaux noethériens.

4.6. Proposition. *Soit A un anneau commutatif. Alors, les propriétés suivantes sont équivalentes :*

(1) *l'anneau A est noethérien ;*

(2) *tout suite croissante d'idéaux de A est stationnaire.*

Autrement dit, si $(\mathfrak{a}_n)_{n\geqslant 0}$ est une suite d'idéaux de A telle que

$$\mathfrak{a}_n \subset \mathfrak{a}_{n+1}, \quad \text{pour tout } n \geqslant 0,$$

alors il existe $N \geqslant 0$ tel que $\mathfrak{a}_n = \mathfrak{a}_N$, pour tout $n \geqslant N$;

(3) *toute famille non vide $\mathscr{F}$ d'idéaux de A admet un élément maximal pour l'inclusion.*

Autrement dit, il existe $\mathfrak{a} \in \mathscr{F}$ tel que pour tout $\mathfrak{a}' \in \mathscr{F}$, on a

$$\mathfrak{a} \subset \mathfrak{a}' \implies \mathfrak{a}' = \mathfrak{a}.$$

Démonstration

(1) $\Longrightarrow$ (2). Supposons que A soit noethérien, et soit $(\mathfrak{a}_n)_{n\geqslant 0}$ une suite croissante d'idéaux de A. Alors, $\mathfrak{a} = \bigcup_{n\geqslant 0} \mathfrak{a}_n$ est un idéal de A. Par hypothèse, il est engendré par un nombre fini d'éléments $a_1, \ldots, a_r \in \mathfrak{a}$. Par définition, pour tout $i \in [\![1, r]\!]$, il existe $n_i \geqslant 0$ tel que $a_i \in \mathfrak{a}_{n_i}$. Soit $N = \max_i n_i$. Alors, $a_i \in \mathfrak{a}_N$, pour tout $i \in [\![1, r]\!]$, et puisque $a_1, \ldots, a_r$ engendrent $\mathfrak{a}$, pour tout $n \geqslant N$, on obtient

$$\mathfrak{a} \subset \mathfrak{a}_N \subset \mathfrak{a}_n \subset \mathfrak{a}.$$

En particulier, on a $\mathfrak{a}_n = \mathfrak{a}_N$, pour tout $n \geqslant N$.

(2) $\Longrightarrow$ (3). Soit $\mathscr{F}$ une famille non vide d'idéaux de A. Supposons que $\mathscr{F}$ n'admette pas d'élément maximal pour l'inclusion, et soit $\mathfrak{a}_0 \in \mathscr{F}$. Puisque l'idéal $\mathfrak{a}_0$ n'est pas un élément maximal pour l'inclusion, il existe $\mathfrak{a}_1 \in \mathscr{F}$ tel que $\mathfrak{a}_0 \subsetneq \mathfrak{a}_1$. De même, il existe $\mathfrak{a}_2 \in \mathscr{F}$ tel que $\mathfrak{a}_1 \subsetneq \mathfrak{a}_2$. Par récurrence, on voit qu'il existe une famille $(\mathfrak{a}_n)_{n\geqslant 0}$ d'éléments de $\mathscr{F}$ telle que $\mathfrak{a}_n \subsetneq \mathfrak{a}_{n+1}$ pour tout $n \geqslant 0$, d'où une contradiction.

(3) $\Longrightarrow$ (1). Soit $\mathfrak{a}$ un idéal de A, et soit $\mathscr{F}$ la famille des idéaux de A de type fini contenus dans $\mathfrak{a}$. Alors, $\mathscr{F}$ est une famille d'idéaux de A non vide, puisqu'elle contient l'idéal nul. Par hypothèse, $\mathscr{F}$ possède un élément maximal $\mathfrak{a}_0$. Par construction, on a $\mathfrak{a}_0 \subset \mathfrak{a}$. Supposons l'inclusion stricte, et soit $a \in \mathfrak{a} \setminus \mathfrak{a}_0$. Alors, l'idéal de A engendré par $\mathfrak{a}_0$ et a est de type fini et contenu dans $\mathfrak{a}$. De plus, il contient strictement $\mathfrak{a}_0$, ce qui contredit la maximalité de $\mathfrak{a}_0$. Ainsi, $\mathfrak{a} = \mathfrak{a}_0$, et $\mathfrak{a}$ est donc de type fini. □

4.7. Remarque. Notons que les chaînes d'idéaux principaux sont fortement liées au problème de factorisation. En effet, si l'on a une chaîne finie

$$(a) = (a_0) \subset (a_1) \subset \cdots \subset (a_n),$$

alors pour tout $i \in [\![1, n]\!]$, il existe $x_i \in A$ tel que $a_{i-1} = a_i x_i$. On obtient alors

$$a = x_1 \cdots x_n a_n.$$

Inversement, une égalité de ce type fournit une chaîne d'idéaux principaux

$$(a) \subset (x_1 \cdots x_n) \subset \cdots \subset (x_1 x_2) \subset (x_1).$$

Remarquons d'ailleurs que le point clef de l'existence d'une décomposition en irréductibles dans un anneau principal est le fait que tout chaîne croissante d'idéaux est stationnaire, i.e. la noethérianité d'un anneau principal.

Dans le cas général, la condition de noethérianité se réinterprète heuristiquement en disant qu'un élément $a \in A \setminus \{0\}$ ne possède qu'un nombre fini de diviseurs. Il n'est donc pas surprenant que les anneaux noethériens aient des propriétés intéressantes vis-à-vis de la factorisation, comme le montre la proposition qui suit.

4.8. Proposition. *Soit A un anneau intègre noethérien. Alors, tout élément $a \in A \setminus \{0\}$ admet une factorisation de la forme*

$$a = u\pi_1 \cdots \pi_r,$$

où $u \in A^\times$ et $\pi_1, \ldots, \pi_r$ sont des éléments irréductibles.

Démonstration. Montrons par l'absurde que tout élément non nul de A est factorisable en produit d'éléments irréductibles et d'un inversible. Pour cela, posons

$$\mathscr{F} = \{(a) \mid a \in A \setminus \{0\},\ a \text{ non factorisable}\}.$$

Supposons que $\mathscr{F}$ soit non vide. La proposition 4.6 implique que $\mathscr{F}$ possède un élément maximal (a_0).

Alors, $a_0 \neq 0$, et a_0 n'est ni inversible, ni irréductible, car sinon il serait factorisable. Il existe donc deux éléments b, $c \in A$ non nuls et non inversibles tels que $a_0 = bc$.

Remarquons aussi que si b et c possédaient une factorisation en produit d'éléments irréductibles et d'un inversible, alors a_0 en aurait une. Ainsi, un de ces deux éléments n'est pas factorisable, disons b. Alors, $(b) \in \mathscr{F}$. Par ailleurs, on a aussi $(a_0) \subset (b)$. Par maximalité de l'idéal (a_0), on obtient alors $(a_0) = (b)$. Puisque l'anneau A est intègre, on en déduit que $a_0 = bu$, avec $u \in A^\times$.

Mais alors, on a

$$bc = a_0 = bu,$$

et par intégrité, on obtient $c = u \in A^\times$, d'où une contradiction. Ainsi, $\mathscr{F}$ est vide, et donc tout élément non nul et non inversible de A se factorise en produit d'un inversible et d'éléments irréductibles. □

4.9. Remarque. On peut démontrer que ce résultat peut être mis en défaut si A est noethérien non intègre (cf. exercice 23).

L'anneau $\mathbb{Z}[i\sqrt{3}]$ est noethérien d'après l'exemple 4.2 (2). En particulier, tout élément non nul se décompose en produit d'irréductibles et d'une unité. En revanche, l'exemple 1.25 montre qu'une telle décomposition n'est pas nécessairement unique à permutation et association des facteurs près. Un anneau dans lequel une telle décomposition existe et est unique est appelé un anneau factoriel, qui est l'objet d'étude du paragraphe suivant.

5. Anneaux factoriels

On va maintenant définir la notion d'anneau factoriel.

5.1. Définition. On dit qu'un anneau A est *factoriel* s'il est intègre, et s'il vérifie les conditions suivantes :

(1) tout élément non nul $a \in A$ peut s'écrire

$$a = u\pi_1 \cdots \pi_r,$$

où $\pi_1, \ldots, \pi_r$ sont irréductibles, et $u \in A^\times$;

(2) la décomposition est unique à permutation et association des facteurs près. Autrement dit, si l'on a deux décompositions du type précédent

$$a = u\pi_1 \cdots \pi_r = u'\pi'_1 \cdots \pi'_s,$$

alors $r = s$ et il existe $\sigma \in \mathfrak{S}_r$ telle que π_i et $\pi'_{\sigma(i)}$ soient associés pour tout $i \in [\![1, r]\!]$.

5.2. Exemple. L'anneau $\mathbb{Z}$ est factoriel. Plus généralement, tout anneau principal est factoriel, d'après le théorème 2.10.

Comme dans le cas d'un anneau principal, le choix d'un s.c.r.i. $\mathscr{P}$ permet d'obtenir une véritable unicité de la décomposition en irréductibles. On peut maintenant reformuler la définition d'un anneau factoriel de la façon suivante.

5.3. Définition. Soit A un anneau, et soit $\mathscr{P}$ un s.c.r.i. Alors, A est factoriel s'il est intègre et si pour tout élément non nul $a \in A$, il existe un unique élément $u \in A^\times$, et une unique famille d'entiers naturels $(n_\pi)_{\pi\in\mathscr{P}}$ tous nuls sauf un nombre fini, tels que

$$a = u \prod_{\pi\in\mathscr{P}} \pi^{n_\pi}.$$

L'équivalence des deux définitions d'un anneau factoriel est laissé au lecteur à titre d'exercice.

Au vu de cette définition, nous pouvons maintenant introduire la notion de valuation par rapport à un élément irréductible.

5.4. Définition. Soit A un anneau factoriel, et soit $\mathscr{P}$ un s.c.r.i. de A.

Soit $a \in A \setminus \{0\}$, et soit

$$a = u \prod_{\pi\in\mathscr{P}} \pi^{n_\pi}$$

sa décomposition en facteurs irréductibles. Pour tout $\pi \in \mathscr{P}$, l'entier n_π est appelé la *valuation π-adique de a*, et est noté $v_p(\pi)$.

On pose aussi $v_\pi(0) = +\infty$ pour tout $\pi \in \mathscr{P}$.

Par définition même, $v_\pi(a)$ est la plus grande puissance[3] de π divisant a. En particulier, $v_\pi(a) \geqslant 1$ si, et seulement si, $\pi \mid a$.

Dans la suite, on utilisera implicitement les conventions

$$+\infty \geqslant m \quad \text{et} \quad (+\infty) + (+\infty) = (+\infty) + m = m + (+\infty) = +\infty,$$

pour tout $m \geqslant 0$.

Les mêmes causes produisant les mêmes effets, on a un analogue parfait du lemme II-4.6.

5.5. Lemme. *Soit A un anneau factoriel, soit $\mathscr{P}$ un s.c.r.i. et soit $\pi \in \mathscr{P}$. Alors :*

(1) *pour tout $a \in A$, on a $v_\pi(a) = +\infty$ si, et seulement si, $a = 0$;*

(2) *pour tous $a, b \in A$, on a $v_\pi(a+b) \geqslant \min\big(v_\pi(a), v_\pi(b)\big)$, avec égalité dès que $v_\pi(a) \neq v_\pi(b)$;*

(3) *pour tous $a, b \in A$, on a $v_\pi(ab) = v_\pi(a) + v_\pi(b)$;*

De plus, pour tous $a, b \in A$, on a $a \mid b$ si, et seulement si, $v_\pi(b) \geqslant v_\pi(a)$ pour tout $\pi \in \mathscr{P}$.

3. Cela est même vrai si $v_\pi(a) = +\infty$, auquel cas $a = 0$, et toute puissance de π divise a, ce que l'on peut traduire symboliquement par la convention $\pi^{+\infty} \mid 0$.

Démonstration. Le premier point provient de la définition. De plus, si $a = 0$ ou $b = 0$, les résultats annoncés sont clairs. On peut donc supposer a et b non nuls. Écrivons

$$a = u \prod_{\pi \in \mathscr{P}} \pi^{v_\pi(a)}, \quad b = u' \prod_{\pi \in \mathscr{P}} \pi^{v_\pi(b)},$$

avec u, $u' \in A^\times$. On a alors

$$ab = uu' \prod_{\pi \in \mathscr{P}} \pi^{v_\pi(a)+v_\pi(b)}.$$

L'unicité de la décomposition montre alors le point (3). Remarquons que l'on a

$$a = \pi^{v_\pi(a)} a' \quad \text{et} \quad b = \pi^{v_\pi(b)} b',$$

où a', $b' \in A$ ne sont pas divisibles par π.

Supposons par exemple que $v_\pi(a) \leqslant v_\pi(b)$. On a alors

$$a + b = \pi^{v_\pi(a)}(a' + \pi^{v_\pi(b)-v_\pi(a)} b').$$

Ainsi, $\pi^{v_\pi(a)} \mid (a+b)$. Or, $v_\pi(a+b)$ est par définition la plus grande puissance de π divisant $a+b$. On en conclut donc que $v_\pi(a+b) \geqslant v_\pi(a)$, ce qu'il fallait vérifier. Supposons maintenant que $v_\pi(a) < v_\pi(b)$. Alors, $\pi \mid \pi^{v_\pi(b)-v_\pi(a)} b'$, et comme $\pi \nmid a'$, on constate que π ne divise pas $a' + \pi^{v_\pi(b)-v_\pi(a)} b'$. D'après le point (3) par exemple, on obtient alors $v_\pi(a + b) = v_\pi(a)$.

Enfin, supposons que $v_\pi(a) \leqslant v_\pi(b)$ pour tout $\pi \in \mathscr{P}$. Alors, on a

$$b = \Big(u \prod_{\pi \in \mathscr{P}} \pi^{v_\pi(a)}\Big)\Big(u'u^{-1} \prod_{\pi \in \mathscr{P}} \pi^{v_\pi(b)-v_\pi(a)}\Big),$$

c'est-à-dire $b = ac$, avec $c = u'u^{-1} \prod_{\pi \in \mathscr{P}} \pi^{v_\pi(b)-v_\pi(a)}$, qui est bien un élément de A, puisque $v_\pi(b) - v_\pi(a) \geqslant 0$ par hypothèse.

Cela achève la démonstration. □

Nous allons maintenant constater que la plupart des propriétés arithmétiques des anneaux principaux sont vraies plus généralement dans les anneaux factoriels. En revanche, les démonstrations dans ce cas sont totalement différentes, et s'appuient (bien sûr) sur l'existence de la décomposition en irréductibles. On commence par le lemme d'Euclide.

5.6. Lemme (Euclide). *Soit A un anneau factoriel. Alors, pour tout élément irréductible π, et tous a, $b \in A$, on a*

$$\pi \mid ab \implies \pi \mid a \text{ ou } \pi \mid b.$$

Autrement dit, tout élément irréductible est premier.

Démonstration. Soit π un élément irréductible de A fixé. Soit $\mathscr{P}$ un s.c.r.i. contenant[4] π, de sorte que l'on peut considérer la valuation π-adique v_π. En vertu du lemme précédent, l'hypothèse de l'énoncé se retraduit alors par $v_\pi(a) + v_\pi(b) \geqslant 1$. Comme $v_\pi(a)$ et $v_\pi(b)$ sont positifs (possiblement égaux à $+\infty$), on a nécessairement $v_\pi(a) \geqslant 1$ ou $v_\pi(b) \geqslant 1$, soit encore $\pi \mid a$ ou $\pi \mid b$.

La dernière partie provient du lemme 1.19. □

5.7. Exemple. Soit $A = \mathbb{Z}[i\sqrt{3}]$. D'après l'exemple 1.25, A n'est pas factoriel. Cela peut aussi se voir en constatant que l'élément irréductible 2 engendre un idéal qui n'est pas premier (cf. exemple 1.23).

On continue par le lemme de Gauss.

5.8. Lemme (Gauss). *Soit A un anneau factoriel, et soient $a, b, c \in A$. Si $a \mid bc$ et si a et b sont premiers entre eux, alors $a \mid c$.*

Démonstration. Fixons un s.c.r.i. $\mathscr{P}$ de A. Par hypothèse, on a $bc = ad$ pour un certain $d \in A$. Pour tout $\pi \in \mathscr{P}$, on a donc

$$v_\pi(b) + v_\pi(c) = v_\pi(a) + v_\pi(d).$$

On va montrer que $v_\pi(c) \geqslant v_\pi(a)$ pour tout $\pi \in \mathscr{P}$, ce qui démontrera le résultat voulu par le lemme 5.5.

Soit $\pi \in \mathscr{P}$. Si $v_\pi(a) = 0$, le résultat voulu est clair. Supposons maintenant que $v_\pi(a) \geqslant 1$, i.e. $\pi \mid a$. Alors, $v_\pi(b) = 0$, car sinon π diviserait b, ce qui n'est pas le cas puisque a et b n'ont pas de diviseurs communs non inversibles. On a alors $v_\pi(c) = v_\pi(a) + v_\pi(d) \geqslant v_\pi(a)$.

Cela achève la démonstration. □

Enfin, on s'intéresse à l'existence d'un pgcd et un ppcm dans un anneau factoriel.

5.9. Lemme. *Soit A un anneau factoriel, et soient $a, b \in A$. Alors, a et b ont un pgcd et un ppcm.*

Démonstration. On vérifie facilement que $\prod_{\pi \in \mathscr{P}} \pi^{\min(v_\pi(a), v_\pi(b))}$ est un pgcd de a et b, et que $\prod_{\pi \in \mathscr{P}} \pi^{\max(v_\pi(a), v_\pi(b))}$ est un ppcm de a et b. Cela est laissé en exercice au lecteur. □

4. Un tel s.c.r.i. existe toujours. En effet, un s.c.r.i. arbitraire contient par définition un élément irréductible π' associé à π. Il suffit de remplacer cet élément par π pour obtenir le s.c.r.i. désiré.

Pour avoir unicité d'un ppcm et d'un pgcd, on introduit la notion d'élément normalisé.

5.10. Définition. Soit A un anneau factoriel, et soit $\mathscr{P}$ un s.c.r.i. On dit qu'un élément $a \in A$ est *normalisé* (par rapport à $\mathscr{P}$) s'il est nul, ou s'il est non nul et s'écrit sous la forme

$$a = \prod_{\pi \in \mathscr{P}} \pi^{n_\pi},$$

avec $n_\pi \geqslant 0$, où les entiers n_π sont tous nuls sauf un nombre fini. Tout élément de A non nul s'écrit alors de manière unique comme produit d'un élément inversible et d'un élément normalisé.

5.11. Remarque. Si l'on fixe un s.c.r.i. d'un anneau factoriel A, tout idéal principal admet alors un unique générateur normalisé. Cela provient du fait que deux générateurs d'un idéal principal dans un anneau intègre diffèrent d'un inversible.

5.12. Exemples

(1) Si $A = \mathbb{Z}$ et si $\mathscr{P}$ est l'ensemble des nombre premiers, un entier est normalisé si, et seulement si, il est positif ou nul.

(2) Si $A = K[X]$ (où K est un corps), et si $\mathscr{P}$ est l'ensemble des polynômes irréductibles unitaires, un polynôme est normalisé si, et seulement si, il est nul ou unitaire.

Comme un pgcd et un ppcm sont uniques à multiplication d'un inversible près, on obtient alors l'unicité d'un unique pgcd normalisé et d'un unique ppcm normalisé, une fois choisi un s.c.r.i.

5.13. Remarques. Soit A un anneau factoriel, soit K_A son corps des fractions et soit $\mathscr{P}$ un s.c.r.i. de A.

(1) Tout élément $x \in K_A^\times$ s'écrit sous la forme $x = \dfrac{a}{b}$, $a \in A$, $b \in A \setminus \{0\}$, avec a et b premiers entre eux.

Il suffit en effet de simplifier par un pgcd d'un numérateur et d'un dénominateur fixés. L'écriture est unique si l'on choisit b normalisé.

(2) Tout élément $x \in K_A^\times$ s'écrit de manière unique sous la forme

$$x = u \prod_{\pi \in \mathscr{P}} \pi^{n_\pi},$$

avec $u \in A^\times$ et $n_\pi \in \mathbb{Z}$, où les entiers n_π sont tous nuls sauf un nombre fini. Cela est aussi un exercice facile, laissé au lecteur.

On continue en donnant quelques caractérisations des anneaux factoriels. On commence par une condition suffisante de factorialité.

5.14. Proposition. *Soit A un anneau intègre. On suppose que tout élément non nul s'écrit comme produit d'un élément inversible et d'éléments premiers. Alors, A est factoriel.*

Démonstration. Puisque A est intègre, un élément premier est irréductible, et l'existence d'une décomposition d'un élément non nul en produit d'un inversible et d'éléments irréductibles est donc assurée.

D'après le lemme 1.26, toute décomposition d'un élément non nul en produit d'une unité et d'éléments irréductibles est égale à association près et permutation près des facteurs à une décomposition fixée en produit d'un élément inversible et d'éléments premiers (qui existe par hypothèse). Cela démontre l'unicité et la factorialité de A. □

On en déduit une première caractérisation des anneaux factoriels.

5.15. Proposition. *Soit A un anneau intègre. Alors, A est factoriel si, et seulement si, les deux conditions suivantes sont vérifiées :*

(1) *tout élément non nul $a \in A$ peut s'écrire*

$$a = u\pi_1 \cdots \pi_r,$$

où $\pi_1, \ldots, \pi_r$ sont irréductibles, et $u \in A^\times$;

(2) *tout élément irréductible est premier.*

Démonstration. Le sens direct provient du lemme d'Euclide, et le sens indirect est donné par la proposition 5.14. □

On peut alors donner une autre démonstration de la factorialité d'un anneau principal.

5.16. Théorème. *Tout anneau principal est factoriel.*

Démonstration. Par définition, un anneau principal est intègre. De plus, tout idéal étant engendré par un élément, il est noethérien. Ainsi, tout élément non nul admet une décomposition en produits d'irréductibles et d'une unité par la proposition 4.8. De plus, tout élément irréductible est premier d'après le lemme 2.5. D'après la proposition 5.15, l'anneau A est factoriel. □

En résumé, on a

$$\text{euclidien} \implies \text{principal} \implies \text{factoriel},$$

mais aucune des implications inverses n'est vraie.

Par exemple, l'anneau $\mathbb{Z}\left[\frac{-1+i\sqrt{19}}{2}\right]$ est principal et non euclidien (cf. exercice 18). On verra aussi dans un chapitre ultérieur que $\mathbb{Z}[X]$ est factoriel. En revanche, il est non principal (cf. exercice XV-36).

Nous continuons ces considérations en énonçant un autre critère de factorialité, et dont la démonstration est plus délicate.

5.17. Théorème. *Soit A un anneau intègre. Alors, A est factoriel si, et seulement si, tout idéal premier non nul contient un élément premier.*

Démonstration. Supposons que A soit un anneau factoriel. Soit $\mathfrak{p}$ un idéal premier non nul. Puisque $\mathfrak{p} \neq A$, $\mathfrak{p}$ contient un élément a non nul, et non inversible. Par hypothèse, on peut écrire

$$a = u\pi_1 \cdots \pi_r,$$

où $u \in A$ et $\pi_1, \ldots, \pi_r$ sont des éléments irréductibles, donc premiers par le lemme d'Euclide. De plus, $r \geqslant 1$, car a est non inversible. Comme $\mathfrak{p}$ est premier et $a \in \mathfrak{p}$, on a $u \in \mathfrak{p}$ ou $\pi_i \in \mathfrak{p}$ pour un certain $i \in [\![1, r]\!]$. Le premier cas est impossible, car $\mathfrak{p} \neq A$. Ainsi, $\mathfrak{p}$ contient un des π_i.

Réciproquement, supposons que tout idéal premier non nul contienne un élément premier. Soit $a \in A$ un élément non nul. Soit S l'ensemble des éléments non nuls de A qui admettent une décomposition en produit d'un inversible et d'éléments premiers. Nous allons démontrer que $S = A \setminus \{0\}$, ce qui établira la factorialité de A par la proposition 5.14.

Par définition, S ne contient pas 0. Montrons maintenant que pour tous éléments $a, b \in A$, on a $ab \in S$ si, et seulement si, $a, b \in S$.

Si $a, b \in S$, il est clair que $ab \in S$, la non nullité de ab découlant de l'intégrité de A.

Pour démontrer l'implication inverse, pour tout $r \geqslant 0$, on note (H_r) la propriété suivante :

(H_r) pour tous $a, b \in A$ tels que $ab = u\pi_1 \cdots \pi_r$, avec $u \in A^\times$ et $\pi_1, \ldots, \pi_r$ premiers, alors $a, b \in S$.

Montrons (H_r) par récurrence sur r. Deux éléments de A dont le produit est inversible sont nécessairement inversibles, donc dans S. Ainsi, (H_0) est vraie.

Supposons que (H_r) soit vraie pour un entier $r \geqslant 0$, et montrons (H_{r+1}). Soient $a, b \in A$ tels que $ab = u\pi_1 \cdots \pi_{r+1}$, avec des notations évidentes. On a donc $\pi_{r+1} \mid ab$, et ainsi $\pi_{r+1} \mid a$ ou $\pi_{r+1} \mid b$ puisque π_{r+1} est premier. Supposons sans perte de généralité que $\pi_{r+1} \mid a$. On a donc $a = \pi_{r+1}a'$,

avec $a' \in A$. Par intégrité, on obtient donc

$$a'b = u\pi_1 \cdots \pi_r.$$

Par hypothèse de récurrence, on en déduit que $a' \in S$ et $b \in S$. Mais alors, on a aussi $a = a'\pi_{r+1} \in S$, puisque S est stable par multiplication. Par conséquent, (H_{r+1}) est vraie, et cela achève la récurrence.

Supposons maintenant qu'il existe un élément a non nul qui n'appartienne pas à S. D'après ce qui précède, (a) est disjoint de S. Le lemme XVI-2.4 montre alors que (a) est contenu dans un idéal premier $\mathfrak{p}$ disjoint de S. Par hypothèse, $\mathfrak{p}$ contient un élément premier. Mais un tel élément premier est dans S, ce qui est impossible par choix de $\mathfrak{p}$. Par conséquent, $S = A \setminus \{0\}$, ce qui achève la démonstration. □

Nous allons maintenant appliquer la notion d'anneau factoriel pour donner une caractérisation des anneaux principaux permettant de réduire l'étude de la principalité des idéaux à celles des idéaux maximaux.

5.18. Théorème. *Soit A un anneau. Alors, A est principal si, et seulement si, il est intègre, noethérien, et tous ses idéaux maximaux sont principaux.*

Démonstration. Si A est principal, alors il est intègre et tout idéal de A (en particulier tout idéal maximal) est principal, donc de type fini. Ainsi, A est noethérien. Réciproquement, supposons que A soit intègre, noethérien, et que tout idéal maximal de A soit principal.

Dans un premier temps, nous allons montrer que A est un anneau factoriel. L'existence d'une décomposition en produit d'éléments irréductibles et d'une unité découle de la proposition 4.8.

Soit $\pi \in A$ un élément irréductible. Il est donc en particulier non inversible, et donc $(\pi) \neq A$. Il existe donc un idéal maximal $\mathfrak{m}$ contenant (π). Par hypothèse, $\mathfrak{m}$ est principal, et donc il existe $a \in A$ tel que $\mathfrak{m} = (a)$. D'après le lemme 2.5, on a $(\pi) = (a) = \mathfrak{m}$. En particulier, (π) est un idéal premier, non nul puisque π est irréductible, et donc π est premier par le lemme 1.19. D'après la proposition 5.15, A est factoriel.

Montrons finalement que A est principal. Soit $\mathfrak{a}$ un idéal de A. Puisque A est noethérien, $\mathfrak{a}$ est de type fini. Ainsi,

$$\mathfrak{a} = (a_1, \ldots, a_n), \quad a_i \in A,\ n \geqslant 1.$$

Si $n = 1$, $\mathfrak{a}$ est principal. Si $n \geqslant 2$, une récurrence facile montre qu'il suffit de considérer le cas $n = 2$.

Supposons que $\mathfrak{a} = (a, b)$. Si $a = 0$ ou $b = 0$, $\mathfrak{a}$ est principal. On peut ainsi supposer que a et b sont tous les deux non nuls. Dans ce cas, puisque A

est factoriel, a et b ont un pgcd, que l'on note d. Remarquons que d est non nul, puisque a et b sont non nuls. Par définition, on a $a, b \in (d)$, d'où l'inclusion $(a,b) \subset (d)$. On peut alors écrire

$$a = d\alpha \quad \text{et} \quad b = d\beta, \quad \text{avec } \alpha, \beta \in A.$$

Alors, α et β sont premiers entre eux. En effet, si $d' \in A$ est un diviseur commun à α et β, dd' est un diviseur commun à a et b. On a alors $dd' \mid d$, soit $d' \mid 1$ puisque d est non nul. Autrement dit, $d' \in A^\times$, ce qu'il fallait vérifier.

On va montrer que $(\alpha, \beta) = A$. Supposons le contraire. Alors, il existe un idéal maximal $\mathfrak{m}$ contenant (α, β). Puisque $\mathfrak{m}$ est principal, on peut alors écrire $\mathfrak{m} = (d')$, avec $d' \in A$. Ainsi, $\alpha, \beta \in (d')$. Il s'ensuit que $d' \mid \alpha$ et $d' \mid \beta$. Comme α et β sont premiers entre eux, $d' \in A^\times$ et donc $\mathfrak{m} = (d') = A$, d'où une contradiction, puisque $\mathfrak{m}$ est maximal. Ainsi, on a $(\alpha, \beta) = A$, et donc il existe $u, v \in A$ tels que $u\alpha + v\beta = 1$. En multipliant par d, on a obtient alors

$$d = ua + vb \in (a,b),$$

et donc $(d) \subset (a,b)$. Ainsi, $(d) = (a,b) = \mathfrak{a}$, et $\mathfrak{a}$ est principal. Cela achève la démonstration. □

6. Exercices

Dans toute la suite, A est un anneau commutatif.

E1.

a. Soit $\pi \in A$ un élément irréductible et soit $u \in A^\times$. Montrer que $u\pi$ est irréductible.

b. Soient $\pi, \pi' \in A$. On suppose que π et π' sont associés. Montrer que π est irréductible si, et seulement si, π' est irréductible.

Attention ! L'anneau A n'est pas supposé intègre.

E2. Montrer que 2 et X sont premiers entre eux dans $\mathbb{Z}[X]$, mais ne sont pas étrangers (cf. définition XVI-3.10).

E3. Dans $\mathbb{R}[X]$, déterminer un pgcd et un ppcm, ainsi qu'un couple de Bézout dans les cas suivants :

(i) $X^3 + X + 4$ et $X^2 + X + 2$;

(ii) $X^4 + 2X^2 - X + 5$ et $X^4 + 3X^2 + X + 1$.

E4. Les anneaux $\mathbb{R}[X]/(X^3 + X + 4)$ et $\mathbb{R}[X]/(X^2 + X + 2)$ sont-ils des corps ? Si oui, les identifier à un corps connu.

E5. Pour tout $n \geqslant 1$, on pose

$$\Delta_n = \operatorname{pgcd}(n^{17} + 9,\ (n+1)^{17} + 9).$$

Le but de cet exercice est d'étudier la suite $(\Delta_n)_{n \geqslant 1}$.

a. À l'aide d'un logiciel de calcul formel, calculer Δ_n pour un très grand nombre de valeurs de n (au moins jusqu'à $n = 10^6$.)

b. Formuler, et essayer de démontrer, une conjecture sur la suite $(\Delta_n)_{n \geqslant 1}$.

c.

(i) Quelles sont les racines complexes de $X^{17} + 9$? Vérifier que $X^{17} + 9$ et $(X+1)^{17} + 9$ n'ont pas de racines complexes communes, et en déduire que les polynômes

$$X^{17} + 9 \quad \text{et} \quad (X+1)^{17} + 9$$

sont premiers entre eux dans $\mathbb{C}[X]$ et dans $\mathbb{Q}[X]$.

(ii) Justifier qu'il existe $A,\ B \in \mathbb{Z}[X]$ et $p \in \mathbb{Z}$ tels que

$$A(X)(X^{17} + 9) + B(X)\big((X+1)^{17} + 9\big) = p.$$

En utilisant un logiciel de calcul formel, expliciter p et vérifier que c'est un nombre premier.

(iii) En déduire que pour tout $n \geqslant 1$, on a $\Delta_n = 1$ ou p.

d. Soit K un corps, soit $a \in K$ et soient $P,\ Q \in K[X]$ deux polynômes non nuls. Soit R_m le dernier reste non nul obtenu en appliquant l'algorithme d'Euclide à P et Q.

Montrer que l'on a

$$\begin{cases} P(a) &= 0 \\ Q(a) &= 0 \end{cases} \iff R_m(a) = 0.$$

e. En utilisant un logiciel de calcul formel, calculer le polynôme R_m lorsque l'on a

$$K = \mathbb{F}_p, \quad P = X^{17} + \overline{9} \in \mathbb{F}_p[X] \quad \text{et} \quad Q = (X + \overline{1})^{17} + \overline{9} \in \mathbb{F}_p[X],$$

où p est le nombre premier défini plus haut.

f. En déduire l'ensemble des entiers $n \geqslant 1$ vérifiant

$$\begin{cases} n^{17} + 9 &\equiv 0 \ [p] \\ (n+1)^{17} + 9 &\equiv 0 \ [p]. \end{cases}$$

g. Quel le plus petit entier $N \geqslant 1$ vérifiant le système de congruences précédent ? Vérifier que $\Delta_n = 1$ pour tout $n < N$. Que vaut Δ_N ? Qu'en conclure quant à la conjecture faite au début de cet exercice ? Moralité ?

E6. Soit $A = \mathbb{Z}[i]$, et soit

$$\begin{aligned} N : A &\longrightarrow \mathbb{N} \\ z &\longmapsto |z|^2. \end{aligned}$$

On rappelle que $N(zz') = N(z) \times N(z')$, pour tous z, $z' \in A$, et que

$$N(z) = 1 \iff z \in A^\times \iff z = \pm 1, \pm i.$$

a. Lister tous les éléments $z \in A$ vérifiant $N(z) \leqslant 10$.

b. Soient z, $z' \in A$. Montrer que si $z \mid z'$ dans A, alors $N(z) \mid N(z')$. Étudier la réciproque.

c. Les éléments 2 et 3 sont-ils irréductibles dans A ?

d. Montrer que si $N(z)$ est un nombre premier, alors z est irréductible dans A. Étudier la réciproque.

e. Les éléments $2 + i$ et $2 - i$ sont-ils irréductibles dans A ?

f. Lister tous les éléments $z \in A$ **irréductibles** tels que $N(z) \leqslant 10$.

g. Décomposer les éléments suivants en produit d'éléments irréductibles dans A : 10, $1 + 5i$, $6 + 8i$, $-13 + 9i$.

h. Sans utiliser l'algorithme d'Euclide, calculer un pgcd de $1+5i$ et $6+8i$.

E7. Calculer un pgcd d de a et b, et un couple de Bézout associé dans les cas suivants :

(i) $A = \mathbb{Z}[i]$, $a = 1 + 5i$, $b = 6 + 8i$;

(ii) $A = \mathbb{Z}[i]$, $a = 5 - i$, $b = 4 + 7i$;

(iii) $A = \mathbb{Z}[j]$, $a = 4 + 2j$, $b = 3$.

E8. Soit $A = \mathbb{Z}[i]$, et soit

$$\begin{aligned} N : A &\longrightarrow \mathbb{N} \\ z &\longmapsto |z|^2. \end{aligned}$$

On s'intéresse à l'ensemble $\Sigma = \{a^2 + b^2 \mid a, b \in \mathbb{Z}\}$.

a. Montrer que $n \in \Sigma$ si, et seulement si, il existe $z \in A$ tel que $n = N(z)$. En déduire que Σ est stable par multiplication.

b. Montrer qu'un nombre premier p n'est pas irréductible dans A si, et seulement si, $p \in \Sigma$.

c. Soit p un nombre premier. Montrer que l'on a des isomorphismes d'anneaux
$$\mathbb{Z}[i]/(p) \simeq \mathbb{Z}[X]/(X^2+1,p) \simeq \mathbb{F}_p[X]/(X^2+\overline{1}).$$

d. En déduire l'équivalence des propriétés suivantes :
(i) p n'est pas irréductible dans A ;
(ii) -1 est un carré modulo p ;
(iii) $p = 2$ ou $p \equiv 1$ [4].

e. Soit $n \geqslant 2$ un entier. Montrer que $n \in \Sigma$ si, et seulement si, $v_p(n)$ est pair pour tout $p \equiv 3$ [4].

Indication. Si $n = a^2+b^2$ et si $p \equiv 3$ [4] divise n, montrer que $p \mid a+ib$ ou $p \mid a-ib$. En déduire que $p^2 \mid n$, et procéder par récurrence.

f. Montrer que les éléments irréductibles de A sont :
(i) les éléments $\pm p$, $\pm ip$, où p est premier, et congru à 3 modulo 4 ;
(ii) les éléments $\pi \in A$ tels que $N(\pi)$ soit un nombre premier.

Indication. Si $\pi \in A$ est irréductible, considérer un diviseur premier p de $N(\pi)$, et distinguer les cas $p \equiv 3$ [4] et $p \equiv 1$ [4]. Dans le second cas, utiliser le fait que $p \in \Sigma$.

g. Montrer que la liste des éléments suivants forment un s.c.r.i. de A :
(i) $1+i$;
(ii) p premier, $p \equiv 3$ [4],
(iii) $a+ib$, où $0 < a < b$ et a^2+b^2 est un nombre premier ;
(iv) $a-ib$, où $0 < a < b$ et a^2+b^2 est un nombre premier.

E9. Soit $P = X^2 - sX + p \in \mathbb{Z}[X]$ un polynôme irréductible. On se fixe une racine $\theta \in \mathbb{C}$ de P, et l'on note $\overline{\theta} \in \mathbb{C}$ l'autre racine de P.

a. Montrer que $\mathbb{Z}[\theta] = \{a+b\theta \mid a, b \in \mathbb{Z}\}$ est un sous-anneau de $\mathbb{C}$, et que tout élément $z \in \mathbb{Z}[\theta]$ s'écrit de manière **unique** sous la forme
$$z = a+b\theta, \quad a, b \in \mathbb{Z}.$$
Pour tout $z = a+b\theta \in \mathbb{Z}[\theta]$, on pose $\overline{z} = a+b\overline{\theta}$.

b. Montrer que pour tout $z \in \mathbb{Z}[\theta]$, on a $\overline{z} \in \mathbb{Z}[\theta]$, et que l'application
$$\begin{aligned} \mathbb{Z}[\theta] &\longrightarrow \mathbb{Z}[\theta] \\ z &\longmapsto \overline{z} \end{aligned}$$
est un morphisme d'anneaux.

c. Vérifier que pour tout $z = a + b\theta \in \mathbb{Z}[\theta]$, on a

$$z + \overline{z} = 2a + sb \quad \text{et} \quad z\overline{z} = a^2 + sab + pb^2.$$

Pour tout $z = a + b\theta \in \mathbb{Z}[\theta]$, on pose

$$N(z) = |z\overline{z}| = |a^2 + sab + pb^2| \in \mathbb{N}.$$

Clairement, d'après ce qui précède, on a

$$N(z_1 z_2) = N(z_1) \times N(z_2), \quad \text{pour tous } z_1,\, z_2 \in \mathbb{Z}[\theta].$$

Le but de ce qui suit est de généraliser certains résultats de l'exercice précédent [5].

d. Soient $z,\, z' \in \mathbb{Z}[\theta]$. Vérifier que si $z \mid z'$ dans $\mathbb{Z}[\theta]$, alors $N(z) \mid N(z')$ dans $\mathbb{Z}$.

e. Soient $a,\, b,\, c \in \mathbb{Z}$. On suppose que a et b sont premiers entre eux dans $\mathbb{Z}$, et que $a + b\theta \mid c$ dans $\mathbb{Z}[\theta]$. Montrer que $N(a + b\theta) \mid c$ dans $\mathbb{Z}$.

Indication. Soient $a,\, a', b,\, b' \in \mathbb{Z}$ tels que $c = (a + b\theta)(a' + b'\theta)$.

Soient $u,\, v \in \mathbb{Z}$ tels que $ua + vb = 1$, et soient

$$M = \begin{pmatrix} a & -sa - pb \\ b & a \end{pmatrix} \quad \text{et} \quad M' = \begin{pmatrix} sb' + a' & v \\ b' & u \end{pmatrix}.$$

Montrer que $MM' = \begin{pmatrix} c & * \\ 0 & 1 \end{pmatrix}$.

Dans toute la suite, on suppose que $\mathbb{Z}[\theta]$ est un anneau principal.

f. Soit $m \in \mathbb{Z}$ et soit $z \in \mathbb{Z}[\theta]$. Soit $\delta \in \mathbb{Z}[\theta]$ un pgcd de m et z dans l'anneau $\mathbb{Z}[\theta]$. Montrer que $\operatorname{pgcd}\big(m, N(z)\big) = N(\delta)$.

Indication. Soit $d = \operatorname{pgcd}\big(m, N(z)\big)$. Utiliser la définition d'un pgcd et les questions précédentes pour démontrer que $N(\delta) \mid d$ dans $\mathbb{Z}$.

Pour démontrer que $d \mid N(\delta)$, justifier que $\delta \in (m, z)$, et montrer qu'il existe $\alpha,\, \beta \in \mathbb{Z}$ tels que $\delta\overline{\delta} = \alpha m + \beta z\overline{z}$.

g. Soit $n \geqslant 0$ un entier et soit $z \in \mathbb{Z}[\theta]$. Soit $\delta \in \mathbb{Z}[\theta]$ un pgcd de n et z. Montrer que si $n \mid N(z)$ dans $\mathbb{Z}$, alors $n = N(\delta)$.

h. Soit $n \geqslant 1$ un entier. En utilisant les questions précédentes, montrer qu'il existe $a,\, b \in \mathbb{Z}$ premiers entre eux tels que $n = N(a + b\theta)$ si, et seulement si, P possède une racine modulo n (i.e. il existe $\omega \in \mathbb{Z}$ tel que $P(\omega) \equiv 0\ [n]$).

Montrer que dans cas, on a $n = N(\delta)$, où δ est un pgcd de n et $\omega - \theta$, où $\omega \in \mathbb{Z}$ est une racine de P modulo n.

5. Les arguments présentés dans les questions suivantes sont de Claude Quitté.

i. Soit p un nombre premier. Montrer qu'il existe des entiers a, $b \in \mathbb{Z}$ tels que $n = N(a+b\theta)$ si, et seulement si, P possède une racine modulo p. Montrer que dans cas, on a $p = N(\delta)$, où δ est un pgcd de p et $\omega - \theta$, où $\omega \in \mathbb{Z}$ est une racine de P modulo p.

j. En déduire les résultats suivants.

(i) Soit $n \geqslant 1$ un entier. Alors, il existe a, $b \in \mathbb{Z}$ premiers entre eux tels que $n = a^2 + b^2$ si, et seulement si, -1 est un carré modulo n.

(ii) Soit p un nombre premier. Alors, il existe des entiers a, $b \in \mathbb{Z}$ tels que $p = a^2 + b^2$ si, et seulement si, $p = 2$ ou $p \equiv 1$ [4].

(iii) Soit p un nombre premier. Alors, il existe des entiers a, $b \in \mathbb{Z}$ tels que $p = a^2 + ab + b^2$ si, et seulement si, $p = 3$ ou $p \equiv 1$ [3].

Expliquer dans chaque cas comment obtenir des entiers a, $b \in \mathbb{Z}$ convenables.

Application numérique

(i) Écrire 409 comme somme de deux carrés. On donne $-1 \equiv 143^2$ [409].

(ii) Écrire 97 sous la forme $a^2 + 2b^2$, a, $b \in \mathbb{Z}$.

(iii) Écrire 157 sous la forme $a^2 + ab + b^2$, a, $b \in \mathbb{Z}$.

E10. Soit $A = \mathbb{Z}[i\sqrt{5}] = \{a + ib\sqrt{5} \mid a, b \in \mathbb{Z}\}$. Soit

$$\begin{aligned} N\colon A &\longrightarrow \mathbb{N} \\ z &\longmapsto |z|^2. \end{aligned}$$

On rappelle que $N(zz') = N(z)N(z')$, pour tous z, $z' \in A$, et que

$$N(z) = 1 \iff z \in A^\times \iff z = \pm 1.$$

a. Résoudre l'équation $N(z) = 9$, $z \in A$.

b. Montrer que les éléments $2 + i\sqrt{5}$, $2 - i\sqrt{5}$ et 3 sont irréductibles, non associés deux à deux, et en déduire deux décompositions distinctes de 9 en produit d'irréductibles.

c. Montrer que l'idéal (3) n'est pas premier.

d. Montrer $A/(3) \simeq \mathbb{F}_3 \times \mathbb{F}_3$. Retrouver le résultat de la question précédente.

Indication. On pourra établir successivement les isomorphismes

$$A/(3) \simeq \mathbb{Z}[X]/(X^2 + 5, 3) \simeq \mathbb{F}_3[X]/(X^2 + \bar{5}).$$

e. Montrer que les diviseurs de 9 dans A sont

$$\pm 1, \quad \pm 3, \quad \pm 9, \quad \pm(2 + i\sqrt{5}) \quad \text{et} \quad \pm(2 - i\sqrt{5}).$$

f. Déterminer l'ensemble des diviseurs communs à 9 et $3(2 + i\sqrt{5})$.

g. En déduire que 9 et $3(2+i\sqrt{5})$ n'ont pas de pgcd.

h. Montrer qu'un éventuel ppcm de 3 et $2+i\sqrt{5}$ dans A divise à la fois 9 et $3(2+i\sqrt{5})$. En déduire que 3 et $2+i\sqrt{5}$ n'ont pas de ppcm.

i. Soit $\mathfrak{a} = (3,\ 1+i\sqrt{5})$.

(i) Montrer que $\mathfrak{a} = \{u+iv\sqrt{5} \mid u \equiv v \mod 3\}$. En déduire que $\mathfrak{a} \neq A$.

(ii) On suppose que $\mathfrak{a} = (\alpha),\ \alpha \in \mathfrak{a}$. Montrer que $\alpha \mid 3$ et $\alpha \mid 1+i\sqrt{5}$. En déduire que $N(\alpha) \mid 3$.

(iii) Montrer que $\mathfrak{a}$ n'est pas un idéal principal.

j. Donner quatre raisons distinctes justifiant le fait que A n'est pas un anneau principal. L'anneau A est-il euclidien pour la fonction N ? pour une autre fonction ?

E11. Dans cet exercice, on note $A = \mathbb{R}[X,Y]/(X^2+Y^2-1)$. On admet que A est un anneau intègre.

a. Montrer que pour tout $P \in \mathbb{R}[X,Y]$, il existe $R, S \in \mathbb{R}[X]$ **uniques** tels que $\overline{P} = \overline{R+SY}$.

b. Montrer que le morphisme d'anneaux

$$\begin{array}{rcl} f : \mathbb{R}[X,Y] & \longrightarrow & A \\ P & \longmapsto & \overline{P(X,-Y)} \end{array}$$

induit un morphisme $\varphi : A \longrightarrow A$, et que φ est un isomorphisme.

c. Montrer que $\overline{R+SY}\varphi(\overline{R+SY}) = \overline{R^2-(1-X^2)S^2}$.
En déduire que pour $a \in A$, il existe un unique polynôme $N(a) \in \mathbb{R}[X]$ tel que $a\varphi(a) = \overline{N(a)}$.

d. Montrer que $N(\overline{1}) = 1$, et que pour tous $a, b \in A$, on a

$$N(ab) = N(a)N(b) \in \mathbb{R}[X].$$

Montrer également que $N(\overline{R+SY}) = R^2-(1-X^2)S^2$.

e. Soit $a \in A$. Montrer que $a = 0$ si et seulement si $N(a) = 0$. Montrer également que $a \in A^\times$ si et seulement si $N(a) \in \mathbb{R}^\times$.

f. Soient $R, S \in \mathbb{R}[X]$. Montrer que $N(\overline{R+SY}) = R^2-(1-X^2)S^2$ est soit nul, soit de degré pair.

On pourra considérer le signe de ce polynôme sur les intervalles $[1,+\infty[$ et $]-\infty,-1]$.

g. Montrer que tout élément $a \in A$ tel que $\deg(N(a)) = 2$ est irréductible dans A.

h. L'anneau A est-il factoriel ?

E12. Soit A un anneau commutatif non trivial. On suppose qu'il existe une fonction $N : A \longrightarrow \mathbb{N}$ vérifiant les conditions suivantes :

(i) pour tout $a \in A$, on a $N(a) = 0$ si, et seulement si, $a = 0$;

(ii) pour tout $a \in A$, on a $a \in A^\times$ si, et seulement si, $N(a) = 1$;

(iii) pour tous $a, b \in A$, on a $N(ab) = N(a)N(b)$.

a. Montrer par récurrence sur $N(a)$ que tout élément $a \in A \setminus \{0\}$ s'écrit sous la forme

$$a = u\pi_1 \cdots \pi_r,$$

où $r \geqslant 0$, $u \in A^\times$, et $\pi_1, \ldots, \pi_r$ sont des éléments irréductibles.

b. On suppose que $A = \mathbb{Z}[\sqrt{d}]$, $\mathbb{Z}\Big[\dfrac{-1+\sqrt{d}}{2}\Big]$ ou $K[X]$ (où K est un corps). Retrouver le fait que tout élément non nul de A possède une décomposition en produit d'un élément inversible et d'éléments irréductibles.

E13. Soit $A = \{P \in \mathbb{Q}[X] \mid P'(0) = 0\}$.

a. Montrer que A est un anneau intègre. Quels sont les éléments inversibles de A ?

b. Montrer que X^2 et X^3 sont des éléments de A. A-t-on $X^2 \mid X^3$ dans l'anneau $\mathbb{Q}[X]$? dans A ?

Dans la suite, tous les idéaux considérés sont des idéaux de A.

c. Montrer que X^2 et X^3 sont irréductibles et non associés dans A. Exhiber un élément de A possédant deux décompositions distinctes en produits d'irréductibles.

d. L'idéal (X^2) est-il premier ?

e. Montrer que $\mathfrak{m} = (X^2, X^3)$ n'est pas un idéal principal.

f. Montrer que les seuls diviseurs unitaires communs à X^5 et X^6 dans A sont 1, X^2, X^3.

g. En déduire que X^5 et X^6 n'ont pas de pgcd.

h. Donner quatre raisons distinctes qui justifient que A n'est pas principal.

i. Vérifier que tout élément de A s'écrit sous la forme

$$P = a_0 + a_3X^3 + X^2Q,$$

avec $Q \in A$ et $a_0, a_3 \in \mathbb{Q}$.

j. Identifier l'anneau $A/\mathfrak{m}$. L'idéal $\mathfrak{m}$ est-il maximal ?

E14. Soit A un anneau. Montrer l'équivalence des trois propriétés suivantes :

(i) A est un corps ;

(ii) $A[X]$ est euclidien ;

(iii) $A[X]$ est principal.

Indication. Pour le sens non trivial, on pourra montrer que (X) est un idéal premier non nul de $A[X]$.

E15. On se propose de résoudre l'équation diophantienne

$$y^2 = x^3 - 2, \quad x, y \in \mathbb{Z}.$$

Soit $A = \mathbb{Z}[i\sqrt{2}]$.

On pose $z = y + i\sqrt{2}$, de sorte que l'équation s'écrit $z\overline{z} = x^3$, $z \in A$.

a. Montrer que x est impair.

b. Montrer que z et $\overline{z}$ sont premiers entre eux.

Indication. On pourra considérer le carré du module d'un diviseur commun, et remarquer que $z - \overline{z} = 2i\sqrt{2}$.

c. En déduire soigneusement que $z = t^3$, avec $t \in A$.

d. Résoudre cette dernière équation, et en déduire toutes les solutions de l'équation initiale.

E16. Soit (A, δ) un anneau euclidien. Montrer qu'il existe $x \in A \setminus A^\times$ tel que l'application

$$\begin{array}{ccc} A^\times \cup \{0\} & \longrightarrow & A/(x) \\ a & \longmapsto & \overline{a} \end{array}$$

obtenue par restriction de la projection canonique est surjective.

Indication. Traiter le cas où A est un corps à part. Si A n'est pas un corps, choisir $x \in A \setminus (A^\times \cup \{0\})$ tel que $\delta(x)$ soit minimal, et établir la surjectivité en utilisant une division euclidienne.

E17. Soit $A = \mathbb{Z}[\alpha]$, avec $\alpha = \dfrac{-1 + i\sqrt{19}}{2}$, de sorte que l'on a

$$A = \{a + b\alpha \mid a, b \in \mathbb{Z}\}.$$

On pose $\nu(z) = |z|^2$ pour tout $z \in A$.

a. Vérifier que $\alpha^2 + \alpha + 5 = 0$, et que

$$\nu(a + b\alpha) = \Big(a - \frac{b}{2}\Big)^2 + \frac{19b^2}{4} = a^2 - ab + 5b^2 \in \mathbb{N}.$$

b. Montrer que $A^\times = \{\pm 1\}$.

On suppose maintenant que A est euclidien pour $\delta : A \setminus \{0\} \longrightarrow \mathbb{N}$.

c. On reprend les notations de l'exercice précédent. Justifier que l'anneau quotient $A/(x)$ admet deux ou trois éléments, puis que $A/(x) \simeq \mathbb{F}_p$, avec $p = 2$ ou 3. On pourra utiliser l'exercice XV-22.

d. En déduire l'existence d'un morphisme d'anneaux $\varphi : A \longrightarrow \mathbb{F}_p$ surjectif, avec $p = 2$ ou 3.

e. Justifier que $\varphi(m) = \overline{m}$ pour tout $m \in \mathbb{Z}$, où $\overline{m}$ est la classe de m modulo p (pour $p = 2$ ou 3).

f. En déduire l'existence d'une racine du polynôme $X^2 + X + \overline{5} \in \mathbb{F}_p[X]$ dans $\mathbb{F}_p$, $p = 2$ ou 3.

g. Montrer que A n'est pas euclidien.

E18. Soit $A = \mathbb{Z}[\alpha]$, avec $\alpha = \dfrac{-1 + i\sqrt{19}}{2}$.

Le but de cet exercice est de montrer que A est principal. En particulier, A est principal, mais non euclidien d'après l'exercice précédent. On rappelle que si $z \in A$, on pose

$$\nu(a + b\alpha) = |a + b\alpha|^2 = a^2 - ab + 5b^2,$$

et que $A^\times = \{\pm 1\}$.

a. Montrer que $A \simeq \mathbb{Z}[X]/(X^2 + X + 5)$, puis les isomorphismes

$$A/(2) \simeq \mathbb{Z}[X]/(2, X^2 + X + 5) \simeq \mathbb{F}_2[X]/(X^2 + X + \overline{1}).$$

En déduire que (2) est un idéal maximal.

b. Soient $z, z' \in A$, $z' \neq 0$. On veut montrer l'existence de $q, r \in A$ tels que

▷ $r = 0$ ou $\nu(r) < \nu(z')$

▷ $z = qz' + r$ ou $2z = qz' + r$.

(i) Montrer que $\dfrac{z}{z'} = u + v\alpha$, pour certains $u, v \in \mathbb{Q}$.

Soit $n \in \mathbb{Z}$ la partie entière de v, de sorte que $n \leqslant v < n + 1$.

(ii) On suppose que $v \notin \left]n + \frac{1}{3}, n + \frac{2}{3}\right]$. Montrer qu'il existe $s, t \in \mathbb{Z}$ tels que

$$|u - s| \leqslant \frac{1}{2}, \quad |v - t| \leqslant \frac{1}{3}.$$

En déduire que $q = s + t\alpha$ et $r = z - qz'$ conviennent.

(iii) On suppose que $v \in \left]n + \frac{1}{3}, n + \frac{2}{3}\right]$. Soit m la partie entière de $2v$. Montrer que

$$2v \notin \left]m + \frac{1}{3}, m + \frac{2}{3}\right],$$

et en déduire l'existence de q et r satisfaisant les conditions requises.

On suppose que $\mathfrak{a}$ est un idéal de A non nul qui n'est pas principal. Soit $z' \in \mathfrak{a} \setminus \{0\}$ tel que $\nu(z')$ soit minimal. Par hypothèse, $\mathfrak{a} \neq (z')$, donc il existe $z \in \mathfrak{a}$ tel que $z \notin (z')$.

c. On suppose que $z = z'q + r$, avec $q, r \in A$, $r = 0$ ou $\nu(r) < \nu(z')$. Montrer que $r \in \mathfrak{a}$. En déduire que $r = 0$, puis une contradiction.

d. On suppose que $2z = z'q + r$, avec $q, r \in A$, $r = 0$ ou $\nu(r) < \nu(z')$.

(i) Montrer que $r = 0$, et donc que $z'q = 2z$.

(ii) Montrer que $z' \in (2)$ ou $q \in (2)$.

(iii) Montrer que le cas $q \in (2)$ ne peut se produire.

On a donc $q \notin (2)$ et $z' \in (2)$. Écrivons $z' = 2w$, $w \in A$, de sorte que $z = wq$.

(iv) Montrer que $(2, q) = A$.

(v) En déduire que $w \in \mathfrak{a}$. Comparer $\nu(w)$ et $\nu(z')$.

e. Conclure.

E19. Soit $A = \mathbb{Z}[X, Y, Z]/\big(X(1 - YZ)\big)$.

a. Montrer que $A^\times = \{\pm\overline{1}\}$.

Indication. Si $P, Q \in \mathbb{Z}[X, Y, Z]$ vérifient $\overline{PQ} = \overline{1}$, on pourra d'abord montrer que $P(X, Y, Y^{-1})Q(X, Y, Y^{-1}) = 1$ dans l'anneau $\mathbb{Q}(Y)[X]$.

b. Trouver deux éléments $a, b \in A$ associés tels que l'on ait $a \neq bu$ pour tout $u \in A^\times$.

Les exercices suivants nécessitent d'avoir étudié l'intégralité de ce chapitre.

E20. Soit A un anneau intègre et soient $a, b \in A$.

a. On suppose que a et b ont un pgcd $d \in A$.

(i) Soit $c \in A$ tel que ca et cb ont un pgcd. Montrer alors que cd est un pgcd de ca et cb.

(ii) On suppose que $(a, b) \neq (0, 0)$, et l'on écrit

$$a = a'd \quad \text{et} \quad b = b'd,$$

avec $a', b' \in A$. Montrer que a' et b' sont premiers entre eux.

On dit qu'un anneau A est un *anneau à pgcd* si A est intègre, et deux éléments quelconques de A ont un pgcd.

On suppose dorénavant que A est un anneau à pgcd.

b. Soient a, b, $c \in A$. Montrer que si a et b sont premiers entre eux, et a et c sont premiers entre eux, alors a et bc sont premiers entre eux.

c. Montrer que le lemme d'Euclide est vrai dans A.

d. Montrer que le lemme de Gauss est vrai dans A.

e. Soit A un anneau dans lequel tout élément non nul admet une factorisation en produit d'un inversible et d'irréductibles. Montrer que les conditions suivantes sont équivalentes :

(1) l'anneau A est factoriel ;

(2) A est un anneau à pgcd.

E21. Soit A un anneau intègre, et soient a, $b \in A$.

a. On suppose que a et b ont un ppcm $m \in A$. Montrer que pour tout élément $r \in A \setminus \{0\}$, rm est un ppcm de ra et rb (on traitera à part les cas $a = 0$ ou $b = 0$).

b. On suppose que a et b ont un ppcm $m \in A$. Justifier qu'il existe $d \in A$ tel que $ab = md$. Montrer que si a et b sont non nuls, alors d est un pgcd de a et b.

c. Soit $r \in A \setminus \{0\}$. On suppose que ra et rb ont un pgcd $d' \in A$. Justifier qu'il existe $d \in A$ tel que $d' = rd$, et montrer que d est un pgcd de a et b.

d. On suppose que pour tout $r \in A \setminus \{0\}$, ra et rb ont un pgcd. Soit d un pgcd de a et b. Justifier qu'il existe $m \in A$ tel que $ab = dm$, et montrer que m est un ppcm de a et b.

e. En déduire que a et b ont un ppcm dans A si, et seulement si, pour tout $r \in A \setminus \{0\}$, ra et rb ont un pgcd.

f. Montrer que dans un anneau intègre, deux éléments quelconques ont un ppcm si, et seulement si, deux éléments quelconques ont un pgcd.

E22. D'après l'exercice 21, si deux éléments fixés d'un anneau intègre ont un ppcm, alors ils ont aussi un pgcd. Le but de cet exercice est de montrer que la réciproque est fausse.

On suppose connus les résultats des exercices 20 et 21.

Soit $d \geqslant 3$ un entier sans facteurs carrés, et soit $A = \mathbb{Z}[i\sqrt{d}]$. On rappelle que pour tout $z \in A$, $|z|^2 \in \mathbb{N}$, et que

$$z \in A^\times \iff |z|^2 = 1 \iff z = \pm 1.$$

a. On suppose que $d + 1 \in \mathbb{Z}$ n'est pas premier, et l'on écrit $d + 1 = pk$, où p est premier et $k \in [\![2, d]\!]$.

(i) Montrer que p est irréductible dans A.

(ii) Montrer que p et $1+i\sqrt{d}$ sont premiers entre eux.

(iii) Montrer que kp et $k(1+i\sqrt{d})$ n'ont pas de pgcd.
Indication. Si $a = kp$ et $b = k(1+i\sqrt{d})$ ont un pgcd, justifier que c'est nécessairement k. Remarquer alors que $1+i\sqrt{d}$ divise a et b.

(iv) En déduire que p et $1+i\sqrt{d}$ n'ont pas de ppcm.

b. On suppose que $d+1$ est un nombre premier, et l'on écrit $d = 2k$, $k \geqslant 1$.

(i) Montrer que 2 est irréductible dans A.

(ii) Montrer que 2 et $2+i\sqrt{d}$ sont premiers entre eux.

(iii) Montrer que $2(k+2)$ et $(k+2)(2+i\sqrt{d})$ n'ont pas de pgcd.

(iv) En déduire que 2 et $2+i\sqrt{d}$ n'ont pas de ppcm.

c. En déduire aussi que $\mathbb{Z}[i\sqrt{d}]$ est non factoriel pour tout $d \geqslant 3$ sans facteurs carrés.

d. En considérant l'élément $p \in A$ si $d+1$ n'est pas un nombre premier, et $2 \in A$ si $d+1$ est un nombre premier, retrouver le résultat de la question précédente.

E23. En considérant l'anneau $\mathbb{Z}/6\mathbb{Z}$, montrer qu'un anneau noethérien non intègre ne possède pas nécessairement d'éléments irréductibles, et que dans un anneau non intègre, un élément premier n'est pas nécessairement irréductible.

En particulier, aucun élément élément non nul et non inversible n'admet de décomposition en produit d'un inversible et d'éléments irréductibles.

E24. On se propose de généraliser les résultats de l'exercice précédent.

a. Soient A_1, A_2 deux anneaux commutatifs non triviaux. Montrer soigneusement que les éléments irréductibles de $A_1 \times A_2$ sont exactement les éléments (π_1, u_2) et (u_1, π_2), avec $u_k \in A_k^\times$ et π_k est un élément irréductible de A_k.

b. Énoncer et démontrer une généralisation du résultat de la question précédente au cas de l'anneau $A_1 \times \cdots \times A_r$, où A_k est un anneau commutatif non trivial.

c. Soit p un nombre premier et soit $r \geqslant 2$.

(i) Justifier que tout élément non nul de $\mathbb{Z}/p^r\mathbb{Z}$ peut s'écrire sous la forme $\overline{m} = \overline{u}\,\overline{p}^k$, où $\overline{u} \in (\mathbb{Z}/p^r\mathbb{Z})^\times$ et $0 \leqslant k \leqslant r-1$.

(ii) Montrer que les éléments irréductibles de $\mathbb{Z}/p^r\mathbb{Z}$ sont exactement les éléments de la forme $\overline{u}\,\overline{p}$, où $\overline{u} \in (\mathbb{Z}/p^r\mathbb{Z})^\times$.

d. Soit $A = \mathbb{Z}/p_1^{r_1}\mathbb{Z} \times \cdots \times \mathbb{Z}/p_s^{r_s}\mathbb{Z}$, où $p_1, \ldots, p_s$ sont des nombres premiers deux à deux distincts, et $r_1, \ldots, r_s$ sont des entiers naturels non nuls. Montrer que les éléments irréductibles de A sont exactement les éléments de la forme $\overline{u}(\overline{p}_i, \cdots, \overline{p}_i)$, avec $u \in A^\times$ et où i est tel que $r_i \geqslant 2$.

Indication. Si $i \neq j$, que dire de $\overline{p}_i \in \mathbb{Z}/p_j^{r_j}\mathbb{Z}$?

e. Montrer que les éléments irréductibles de $\mathbb{Z}/n\mathbb{Z}$ sont exactement les éléments de la forme $\overline{u}\,\overline{p}$, où $\overline{u} \in (\mathbb{Z}/n\mathbb{Z})^\times$ et p est un nombre premier tel que $p^2 \mid n$.

f. Soit $n \geqslant 2$.

(i) Justifier que les idéaux de $\mathbb{Z}/n\mathbb{Z}$ sont principaux.

(ii) Déterminer les idéaux premiers de $\mathbb{Z}/n\mathbb{Z}$. On pourra utiliser un des théorèmes d'isomorphisme.

(iii) Montrer que les éléments premiers de $\mathbb{Z}/n\mathbb{Z}$ sont exactement les éléments de la forme $\overline{u}\,\overline{p}$, où $\overline{u} \in (\mathbb{Z}/n\mathbb{Z})^\times$ et p est un nombre premier tel que $p \mid n$, sauf dans un cas particulier que l'on précisera.

g. Soit $n \geqslant 2$.

(i) Énoncer et démontrer une condition nécessaire et suffisante sur n pour que $\mathbb{Z}/n\mathbb{Z}$ n'ait aucun élément irréductible.

(ii) Énoncer et démontrer une condition nécessaire et suffisante sur n pour que $\mathbb{Z}/n\mathbb{Z}$ n'ait aucun élément premier.

(iii) Un élément premier de $\mathbb{Z}/n\mathbb{Z}$ (lorsqu'il existe) est-il nécessairement irréductible ?

(iv) On suppose que $\mathbb{Z}/n\mathbb{Z}$ possède des éléments irréductibles. Montrer que $\overline{m}$ est irréductible si, et seulement si, il est premier.

E25. Démontrer la proposition 5.9.

E26. Soit A un anneau factoriel, et soit $\mathscr{P}$ un s.c.r.i.

a. Montrer que tout $x \in K_A$ non nul s'écrit de manière unique sous la forme

$$x = u \prod_{\pi \in \mathscr{P}} \pi^{n_\pi},$$

où $u \in A^\times$, $n_\pi \in \mathbb{Z}$, et les entiers n_π sont presque tous nuls.

L'entier $n_\pi \in \mathbb{Z}$ est encore noté $v_\pi(x)$. On pose aussi $v_\pi(0) = +\infty$.

b. Soit $\pi \in \mathscr{P}$. Montrer les propriétés suivantes :

(i) pour tout $x \in K_A$, on a $v_\pi(x) = \infty$ si, et seulement si, $x = 0$;

(ii) pour tous $x, y \in K_A$, on a $v_\pi(x + y) \geqslant \min\big(v_\pi(x), v_\pi(y)\big)$, avec égalité dès que $v_\pi(x) \neq v_\pi(y)$;

(iii) pour tous $x, y \in K_A$, on a $v_\pi(xy) = v_\pi(x) + v_\pi(y)$.

(iv) pour tout $x \in K_A$, on a $x \in A$ si, et seulement si, $v_\pi(x) \geqslant 0$ pour tout $\pi \in \mathscr{P}$.

E27. On dit qu'un anneau A est un *anneau de Bézout* si A est intègre, et tout idéal de type fini est principal.

a. Montrer qu'un anneau de Bézout est un anneau à pgcd (i.e. deux éléments quelconques ont un pgcd).

b. Soit A un anneau de Bézout. Montrer que les conditions suivantes sont équivalentes :

(i) A est noethérien ;

(ii) A est un anneau principal ;

(iii) A est factoriel.

Indication. Si $\mathfrak{a}$ est un idéal non de type fini, soit $a \in \mathfrak{a}$ non nul et non inversible. Montrer qu'il existe une suite $(a_n)_{n\geqslant 1}$ d'éléments de A tels que

$$(a) \subsetneq (a_1) \subsetneq \cdots \subsetneq (a_n) \subsetneq \cdots$$

En déduire que a possède une infinité de diviseurs non associés deux à deux, et en déduire que A n'est pas factoriel.

E28. Le but de cet exercice est d'exhiber un anneau de Bézout (et donc un anneau à pgcd) non factoriel.

Soit $A = \{P \in \mathbb{Q}[X] \mid P(0) \in \mathbb{Z}\}$.

a. Montrer que A est un anneau intègre et $A^\times = \{\pm 1\}$.

b. Montrer que les éléments irréductibles de A sont :

(1) $\pm p$, p nombre premier ;

(2) $\pm P$, où $P \in \mathbb{Q}[X]$ est irréductible dans $\mathbb{Q}[X]$ et vérifie $P(0) = 1$.

On pourra remarquer, en le justifiant, que si $P(0) = 0$, alors $m \mid P$, pour tout $m \in \mathbb{Z}$ non nul.

c. Montrer que X n'a pas de décomposition en produit d'un inversible et d'éléments irréductibles. En déduire que A n'est pas factoriel.

d. Quels sont les éléments qui admettent une factorisation en produit d'irréductibles ? La factorisation est-elle unique dans ce cas ?

e. Montrer que A n'est pas noethérien.

f. Montrer que l'idéal $\left(\dfrac{X}{2^n},\ n \geqslant 1\right)$ n'est pas principal.

g. On veut démontrer que A est un anneau de Bézout. Clairement, il suffit de montrer que pour tous P, $Q \in A$, l'idéal (P, Q) est principal. On peut évidemment supposer que P et Q sont non nuls.

(i) Vérifier qu'il suffit de montrer qu'il existe $D \in (P, Q)$ tels que $D \mid P$ et $D \mid Q$.

(ii) Vérifier que tout $R \in \mathbb{Q}[X]$ non nul s'écrit de manière unique

$$R = cX^m R_0,$$

où $c \in \mathbb{Q}^\times$, $m \geqslant 0$, et où $R_0 \in \mathbb{Q}[X]$ vérifie $R_0(0) = 1$.

(iii) Montrer que si $P(0) = Q(0) = 0$, on peut écrire

$$(P, Q) = cX^d(P_1, Q_1),$$

avec P_1, $Q_1 \in A$, et $\big(P_1(0), Q_1(0)\big) \neq (0, 0)$. En déduire que l'on peut supposer que $Q(0) \neq 0$, puis que P et Q sont premiers entre eux dans $\mathbb{Q}[X]$.

On suppose maintenant que P et Q sont premiers entre eux.

(iv) Montrer qu'il existe deux polynômes U_0, $V_0 \in A[X]$ et $r \in \mathbb{Z}$ non nul tels que $U_0 P + V_0 Q = r$.

(v) Soit d un pgcd de $P(0)$ et $Q(0)$. Montrer que d est un diviseur commun à $P - P(0)$ et $Q - Q(0)$ dans A, puis que d est un diviseur commun à P et Q dans A.

(vi) Montrer qu'il existe a, $b \in \mathbb{Z}$ tels que $(aP + bQ - d)(0) = 0$. En déduire que $aP + bQ - d \in rA$, puis que $d \in (P, Q)$ et conclure.

E29. Soit $\mathcal{H}$ l'anneau des fonctions entières (i.e. holomorphes sur $\mathbb{C}$).

a. En utilisant le principe du prolongement analytique, montrer que $\mathcal{H}$ est un anneau intègre.

b. Montrer que les éléments inversibles de $\mathcal{H}$ sont les fonctions entières ne s'annulant pas.

c. Montrer que pour tout $z_0 \in \mathbb{C}$, $z - z_0$ est irréductible. Réciproquement, montrer que les fonctions $z - z_0$, $z_0 \in \mathbb{C}$ sont les seuls éléments irréductibles de $\mathcal{H}$, à association près.

Indication. Une fonction irréductible n'est pas inversible, donc s'annule en un point.

d. En déduire que $f \in \mathcal{H}$ possède une factorisation en produit d'irréductibles et d'un inversible si, et seulement si, la fonction entière f possède un nombre fini de zéros.

e. Donner un exemple d'une fonction $f \in \mathcal{H}$ qui ne possède pas de factorisation, et en déduire que $\mathcal{H}$ n'est pas noethérien.

f. En considérant l'idéal engendré par les fonctions

$$\begin{array}{rccc} f_n : & \mathbb{C} & \longrightarrow & \mathbb{C} \\ & z & \longmapsto & \dfrac{\sin(\pi z)}{z\times(z-1)\times\cdots\times(z-n)} \end{array},$$

montrer directement que $\mathcal{H}$ n'est pas noethérien.

E30.

a. Soient e et e' deux idempotents d'un anneau commutatif A. Montrer que les propriétés suivantes sont équivalentes :

(1) $e' \mid e$;

(2) $e = ee'$;

(3) $(e) \subset (e')$;

(4) $(1-e') \subset (1-e)$.

En déduire que si $(e) = (e')$, alors $e = e'$.

Indication. Si $ee' = e$, on remarquera que $1 - e' = (1-e)(1-e')$.

b. On suppose que A est noethérien. Montrer que toute famille d'idéaux principaux engendrés par un idempotent possède un élément minimal.

On dit qu'un idempotent $e \in A$ est *primitif* s'il ne s'écrit pas comme la somme de deux idempotents orthogonaux non nuls.

c. Montrer qu'un idempotent $e \in A$ est primitif si, et seulement si, eA est un anneau sans idempotents non triviaux.

Indication. Si e est primitif, et si $x \in eA$ est un idempotent, montrer que x et $e - x$ sont des idempotents orthogonaux. Si eA est sans idempotents non triviaux, et si $e = e_1 + e_2$, où e_1, e_2 sont des idempotents orthogonaux, remarquer que $e_1 = e_1 e$ et $e_2 = e_2 e$.

d. On suppose que A est noethérien. Montrer que tout idempotent de A est une somme finie d'idempotents primitifs deux à deux orthogonaux.

Indication. Procéder par l'absurde, et considérer un idempotent $e \in A$ qui n'est pas somme finie d'idempotents primitifs deux à deux orthogonaux tel que (e) soit minimal. Montrer que e est non nul, et qu'il existe deux idempotents e_1, e_2 orthogonaux non nuls tels que $e = e_1 + e_2$. Montrer que $e_1 = e_1 e$ et $e_2 = e_2 e$, puis que $e_1 = e$ ou $e_2 = e$, et aboutir à une contradiction.

e. En déduire que tout anneau noethérien est isomorphe à un produit direct d'un nombre fini d'anneaux noethériens sans idempotents non triviaux.

E31. Soit A un anneau non trivial.

a. On suppose que A n'a pas d'éléments idempotents non triviaux. Montrer que $\pi \in A\setminus\{0\}$ est irréductible si, et seulement si, (π) est maximal parmi les idéaux principaux distincts de A.

En déduire que si de plus A est noethérien et n'est pas un corps, alors A possède un élément irréductible.

Indication. Si (π) est maximal parmi les idéaux principaux distincts de A, vérifier que π est non inversible. Si $\pi = ab$, avec a et b non inversibles, montrer qu'il existe c, $d \in A$ tels que $a = c\pi$ et $b = d\pi$. Montrer alors que πcd est un idempotent de A.

b. On suppose que $A = A_1 \times \cdots \times A_n$, où chaque A_i est noethérien non trivial, sans idempotents non triviaux.

(i) S'il existe un anneau A_i qui n'est pas un corps, déduire de la question précédente que A possède un élément irréductible.

(ii) Si tous les anneaux A_i sont des corps et $n \geqslant 2$, montrer que A possède un élément premier.

c. En utilisant l'exercice 30, montrer que tout anneau noethérien non trivial qui n'est pas un corps possède un élément qui est soit irréductible, soit premier [6].

E32. On dit que $\pi \in A$ est *indivisible* s'il est non nul, non inversible, et si pour tous a, $b \in A$, on a

$$\pi = ab \quad \Longrightarrow \quad \pi \mid a \text{ ou } \pi \mid b.$$

a. Montrer qu'un élément premier est indivisible.

b. Montrer qu'un élément irréductible est indivisible. Montrer que si A est intègre, un élément est indivisible si, et seulement si, il est irréductible.

c. Montrer que dans un anneau noethérien, tout élément s'écrit comme produit d'un élément inversible et d'éléments indivisibles.

E33. Soit A un anneau commutatif fini. On veut montrer que tout élément indivisible [7] est irréductible ou premier [8].

Pour cela, on procède par l'absurde. Soit $\pi \in A$ un élément indivisible, non irréductible et non premier.

a. Montrer qu'il existe $d \in A$ non nul et non inversible tel que $\pi = \pi d$.

Indication. Utiliser le fait que π est indivisible, et non irréductible.

6. Ce résultat m'a été signalé par YCor, un utilisateur très actif de MathOverflow.
7. La notion d'élément indivisible est définie dans l'exercice 32.
8. Ce résultat m'a été proposé par Pace Nielsen.

b. Justifier que d est un diviseur de zéro. Soit $c \in A \setminus \{0\}$ tel que $cd = 0$. En considérant $(\pi + c)d$, montrer que $\pi \mid c$ ou $\pi \mid d$.

c. On suppose que $\pi \mid c$. Aboutir à la contradiction $c = 0$.

d. On suppose que $\pi \mid d$. Soit $a \in A$ tel que $d = \pi a$.

(i) Montrer que d est un idempotent non trivial, et que $(\pi) = (d)$.

(ii) Justifier que l'on a un isomorphisme d'anneaux

$$A/(\pi) \simeq (1-d)A,$$

et en déduire que $(1-d)A$ est non trivial, et non intègre.

(iii) Soient s, $t \in (1-d)A$ non nuls tels que $st = 0$. Montrer que l'on a alors $\pi = (\pi + s)(d + t)$, mais que $\pi \nmid (\pi + s)$ et $\pi \nmid (d + t)$.

e. Conclure.

E34. Soit K un corps et soit $A = K[X,Y]/\big(X^2, X(Y-1)\big)$.

a. Montrer que tout élément de A est représenté par un élément de la forme

$$aX + P(Y), \quad a \in K,\ P \in K[Y],$$

où a et P sont uniques.

En déduire que le morphisme évident $K[Y] \longrightarrow A$ est injectif.

b. Montrer que pour tous a, $b \in K$, et tous P, $Q \in K[Y]$, on a

$$\overline{aX + P(Y)}\ \overline{bX + Q(Y)} = \overline{\big(aQ(1) + bP(1)\big)X + P(Y)Q(Y)}.$$

c. En déduire que $A^\times = \{\overline{aX + \alpha} \mid a \in K,\ \alpha \in K^\times\}$.

d. Montrer que A est non trivial, et non intègre.

e. Montrer que les éléments irréductibles de A sont les éléments

$$\overline{\pi(Y)}, \quad \overline{aX + \pi(Y)}, \quad \overline{aX + b(Y-1)^n}$$

où a, $b \in K^\times$ et $\pi \in K[Y]$ est irréductible.

f. Montrer que $A/(\overline{X}) \simeq K[X,Y]/\big(X^2, X, X(Y-1)\big) \simeq K[Y]$. En déduire que pour tout $a \in K^\times$, $\overline{aX}$ est premier. Montrer plus précisément que si $\overline{R}$, $\overline{S} \in A$ vérifient $\overline{R}\ \overline{S} = \overline{X}$, alors on a $\overline{R} = \overline{rX}$ ou $\overline{S} = \overline{sX}$, avec r, $s \in K^\times$.

g. Soit $\pi \in K[Y]$.

(i) Montrer que $A/(\overline{\pi})$ contient un sous-anneau isomorphe à $K[Y]/(\pi)$. En déduire que si $\overline{\pi}$ est un élément premier, π est irréductible dans l'anneau $K[Y]$.

(ii) On suppose que $\pi(1) = 0$. Vérifier que $\pi = a(Y-1)$, $a \in K^\times$. Montrer que $A/(\overline{\pi}) \simeq K[X]/(X^2)$, et en déduire que $\overline{\pi}$ n'est pas premier.

(iii) On suppose que $\pi(1) \neq 0$. Montrer que $Y-1$ est alors inversible modulo $\pi(Y)$, et en déduire les isomorphismes

$$A/(\overline{\pi}) \simeq K[X,Y]/(X, \pi(Y)) \simeq K[Y]/(\pi).$$

En déduire que $\overline{\pi}$ est premier.

h. Soit $P \in K[Y]$ et soit $a \in K^\times$.

(i) Montrer que $A/(\overline{aX + P(Y)}) \simeq K[Y]/\big(P(Y)^2, P(Y)(Y-1)\big)$.

(ii) Montrer que si $P(1) \neq 0$, $P \in \big(P(Y)^2,\ P(Y)(Y-1)\big)$. En déduire alors que $A/(\overline{aX + P(Y)}) \simeq K[Y]/(P)$, puis que $\overline{aX + P(Y)}$ est premier si, et seulement si, P est irréductible.

(iii) Si $P = (Y-1)Q$, $Q \in K[Y]$, montrer que $\overline{aX + P(Y)}$ n'est pas premier.

i. En déduire que les éléments premiers de A sont les éléments

$$\overline{\pi(Y)}, \quad \overline{aX + \pi(Y)}, \quad \overline{aX},$$

où $a \in K^\times$ et $\pi \in K[Y]$ est un polynôme irréductible tel que $\pi(1) \neq 0$.

E35. Soit $A = \mathbb{F}_2[X,Y,Z]/\big(X^2 - Y^2,\ X(Z-1),\ Y(Z-1)\big)$. On note x, y, z les classes de X, Y, Z dans A. On supposera connus les résultats de l'exercice précédent.

a. Montrer que tout élément $a \in A$ s'écrit sous la forme

$$a = yP(y) + xQ(y) + R(z),$$

où P, Q, $R \in \mathbb{F}_2[X,Y,Z]$ sont uniques.

b. Montrer que x, y, z sont non nuls.

c. En considérant l'anneau quotient $A/(z)$, montrer que z est non inversible. En déduire que x et y ne sont pas irréductibles.

Le but des questions suivantes est de montrer que x est indivisible.

d. Montrer qu'il existe un unique isomorphisme d'anneaux

$$\varphi : A/(y) \xrightarrow{\sim} \mathbb{F}_2[X,Z]/\big(X^2,\ X(Z-1)\big)$$

qui envoie x, y, z sur les classes de X, 0 et Z modulo $\big(X^2,\ X(Z-1)\big)$ respectivement.

e. Soient a, $b \in A$ tels que $ab = x$. On écrit $a = yP(y) + xQ(y) + R(z)$. En utilisant les résultats de l'exercice précédent, montrer que l'on peut

supposer par exemple que

$$XQ(0) + R(Z) \equiv X \mod \big(X^2,\ X(Z-1)\big).$$

Montrer que $Q(0) = 1$, et que $R = 0$.

f. On écrit $b = yS(y) + xT(y) + U(z)$. Montrer que dans l'anneau $\mathbb{F}_2[Y]$, on a les égalités

$$Y(PS + QT) + U(\overline{1})P = \overline{0} \quad \text{et} \quad (PT + QS)Y + U(\overline{1})Q = \overline{1}.$$

g. En déduire successivement les égalités :

(i) $U(\overline{1}) = \overline{1}$;

(ii) $(YS + \overline{1})P + YTQ = \overline{0}$ et $YTP + (YS + \overline{1})Q = \overline{1}$;

(iii) $\big(Y(S+T) + \overline{1}\big)^2 P = YT$ et $\big(Y(S+T) + \overline{1}\big)^2 Q = YS + \overline{1}$;

(iv) $\big(Y(S+T) + \overline{1}\big)(P+Q) = \overline{1}$;

(v) $S = T$, $Q = P + \overline{1}$ et $P = YS$.

h. En déduire que $x \mid a$, et en conclure que x est un élément indivisible, non irréductible et non premier[9].

9. Cet exemple m'a été proposé par Pace Nielsen.

Chapitre XIX

Anneaux à division et codes Wifi

Depuis quelques années, la théorie des anneaux à division a trouvé de très belles applications en théorie du codage, et plus spécifiquement dans le codage Wifi. Devant la recrudescence des téléphones portables/smartphones et autres bêtes nuisibles du même genre, et puisque le pluridisciplinaire est de plus en plus à la mode, l'occasion était trop belle d'écrire un chapitre pour vous expliquer comment ça marche [1] !

Nous avons ici pris le parti d'écrire un texte pouvant être lu par un étudiant de L3. Aussi, les raisons profondes pour lesquelles tout ce qui suit fonctionne ont souvent été cachées. Le lecteur qui serait intéressé par la théorie générale des anneaux à division et son application au codage Wifi, pourra consulter [12], qui fournit toutes les explications nécessaires qui manqueraient ici, en n'omettant aucun détail. Attention néanmoins, cet ouvrage, rédigé dans la langue de Shakespeare, est plutôt destiné à un bon étudiant de M1 ayant des bases solides en algèbre.

Mais revenons au problème qui nous préoccupe, et commençons par expliquer le contexte.

1. Le problème du codage

Personne ne contestera aujourd'hui le fait qu'il est vital de pouvoir transmettre une grande quantité d'information en un minimum de temps. C'est pourquoi, en particulier pour la transmission sans fil, on utilise des codes multi-antennes. Autrement dit, l'émetteur et le récepteur disposent chacun de plusieurs antennes. Les signaux envoyés seront des nombres complexes,

1. Le contenu de ce chapitre est une version remaniée de [11].

contrairement au cas des codes linéaires auto-correcteurs (dont les mots de code sont des vecteurs à valeurs dans un corps fini, comme $\mathbb{F}_2$ par exemple).

Au cours d'une utilisation, chaque antenne émettrice va envoyer un signal différent, et chaque signal, du fait de l'environnement sans fil, va suivre différents parcours avant d'être réceptionné par chaque antenne réceptrice. En fait, au cours de la transmission sans fil, il va se produire deux phénomènes : une atténuation de la force du signal, qui se produit par exemple lorsque le signal rencontre des obstacles, et l'addition de bruit [2]. Ainsi, chaque antenne réceptrice ne va pas recevoir un seul signal, mais une somme de signaux : les signaux envoyés par chaque antenne émettrice, affectés d'un coefficient d'atténuation, plus du bruit.

Rentrons maintenant un peu plus dans les détails. Supposons que l'on dispose de deux antennes émettrices et de deux antennes réceptrices.

Au cours de la première utilisation, l'antenne émettrice i va envoyer un signal, que l'on notera x_{i1}. Comme expliqué précédemment, l'antenne réceptrice j va recevoir un signal composé de x_{11} et x_{21}, affectés d'un coefficient d'atténuation et de bruit v_{j1}. Autrement dit, on a

$$\begin{aligned} y_{11} &= h_{11}x_{11} + h_{12}x_{21} + v_{11} \\ y_{21} &= h_{21}x_{11} + h_{22}x_{21} + v_{21} \end{aligned}$$

On fait maintenant une hypothèse, à savoir qu'il existe un intervalle de temps (très court) au cours duquel les propriétés du canal de transmission sont constantes, ce qui est somme toute raisonnable : si un arbre ou un immeuble est là au temps $t = 1$, il sera encore là quelques milli-secondes après [3], et un véhicule en mouvement ne se déplace pas tant que ça en un laps de temps aussi court. On traduit cela par le fait que les coefficients d'atténuation ne changeront pas au cours de la seconde utilisation.

Ainsi, lors de la seconde transmission, l'antenne émettrice i enverra un signal x_{i2} et les signaux reçus à l'autre bout seront

$$\begin{aligned} y_{12} &= h_{11}x_{12} + h_{12}x_{22} + v_{12} \\ y_{22} &= h_{21}x_{12} + h_{22}x_{22} + v_{22} \end{aligned}$$

En résumé : on a $\mathbf{Y} = \mathbf{HX} + \mathbf{V} \in \mathrm{M}_2(\mathbb{C})$, avec des notations évidentes.

Bien entendu, on obtient une équation similaire si l'on suppose que l'on dispose de n antennes émettrices/réceptrices. Ainsi, les mots de code envoyés sont des matrices $\mathbf{X}$, qui appartiendront à un sous-ensemble fini (de grand cardinal) $\mathcal{C} \subset \mathrm{M}_n(\mathbb{C})$. En revanche, il n'y a aucune raison a priori que la matrice $\mathbf{Y}$ reçue soit encore un mot de code. Il s'agit donc de trouver une manière de récupérer $\mathbf{X}$ à partir de $\mathbf{Y}$.

2. La fameuse « friture » sur la ligne.
3. Sauf si Godzilla est passé par là entre-temps.

Dans ce chapitre, on va supposer que le récepteur a une connaissance parfaite du *code* $\mathcal{C}$ (i.e. l'ensemble des matrices $\mathbf{X}$ qui seront envoyées), et des propriétés du canal de transmission, c'est-à-dire qu'il connaît la matrice $\mathbf{H}$. Dans ce cas, pour décoder, on procède comme suit : ayant reçu $\mathbf{Y}$, on calcule toutes les matrices $\mathbf{HX}, \mathbf{X} \in \mathcal{C}$, et l'on choisit (grâce à un algorithme approprié) un mot de code $\hat{\mathbf{X}} \in \mathcal{C}$ qui minimise $||\mathbf{H}\hat{\mathbf{X}} - \mathbf{Y}||$, où la norme d'une matrice $M = (m_{ij}) \in \mathrm{M}_n(\mathbb{C})$ est définie ici par

$$||M|| = \sqrt{\sum_{i,j} |m_{ij}|^2}.$$

Bien entendu, notre intérêt est que l'information transmise soit récupérée à la sortie, c'est-à-dire que $\hat{\mathbf{X}} = \mathbf{X}$, et donc de minimiser la probabilité de se tromper, probabilité que l'on notera $\mathbb{P}(\mathbf{X} \to \hat{\mathbf{X}} \neq \mathbf{X})$.

On montre alors que, pour tous $\mathbf{X}$, $\hat{\mathbf{X}} \in \mathcal{C}$, on a

$$\mathbb{P}(\mathbf{X} \to \hat{\mathbf{X}} \neq \mathbf{X}) \leqslant \frac{\kappa}{\delta_{\min}(\mathcal{C})^n},$$

où n est le nombre d'antennes émettrices/réceptrices, et où

$$\delta_{\min}(\mathcal{C}) = \inf\{|\det(\mathbf{X}_1 - \mathbf{X}_2)|^2 \mid \mathbf{X}_1 \neq \mathbf{X}_2 \in \mathcal{C}\}.$$

1.1. Définition. La quantité $\delta_{\min}(\mathcal{C})$ s'appelle le *déterminant minimum* du code $\mathcal{C}$.

La quantité κ, quant à elle, est une fonction du déterminant minimum et du *rapport signal-bruit* ρ. Ce rapport signal-bruit, que nous ne définirons pas de manière précise, est le rapport entre la valeur moyenne de l'énergie utilisée pour envoyer $\mathbf{X}$, qui est mesurée par $||\mathbf{X}||^2$, et l'amplitude du bruit.

Plus ρ est grand, plus on utilise de puissance pour couvrir le bruit, et meilleure sera la transmission..., mais cela coûtera plus cher ! La fonction κ possède les propriétés suivantes : elle décroît vers 0 quand ρ tend vers l'infini, et ce d'autant plus vite que $\delta_{\min}(\mathcal{C})$ est grand !

Conclusion. Au vu de la borne ci-dessus, un grand déterminant minimum assure donc une probabilité d'erreur petite, sans demander trop de puissance pour couvrir le bruit.

Mais attention, la solution consistant à multiplier les éléments de $\mathcal{C}$ par un scalaire ne convient pas, puisque cela revient à augmenter le rapport signal-bruit, donc la quantité d'énergie requise pour la transmission. Or, l'énergie coûte cher, et la quantité d'énergie disponible est fixée à l'avance.

Il s'agit donc d'être plus malin que cela, et choisir notre code $\mathcal{C}$ de manière adéquate. La première étape est déjà de s'assurer que $\delta_{\min}(\mathcal{C}) \neq 0$!

Ce simple problème n'est déjà pas complètement trivial, à cause de la non-linéarité du déterminant. Même si l'on s'arrange pour que $\mathcal{C}$ ne soit constitué que de matrices inversibles, cela ne garantit pas pour autant que la différence de deux éléments de $\mathcal{C}$ distincts le soit !

C'est là que les anneaux à division interviennent de manière naturelle.

1.2. Lemme. *Soit D un sous-anneau de $\mathrm{M}_n(\mathbb{C})$ qui soit aussi un anneau à division. Alors, tout sous-ensemble fini $\mathcal{C} \subset D$ vérifie $\delta_{\min}(\mathcal{C}) > 0$.*

Démonstration. Soient $\mathbf{X}_1, \mathbf{X}_2 \in \mathcal{C} \subset D, \mathbf{X}_1 \neq \mathbf{X}_2$. On ne peut raisonnablement pas espérer que $\mathbf{X}_1 - \mathbf{X}_2 \in \mathcal{C}$, mais on a quand même $\mathbf{X}_1 - \mathbf{X}_2 \in D$, puisque D est un sous-anneau de $\mathrm{M}_n(\mathbb{C})$. Par hypothèse, $d = \mathbf{X}_1 - \mathbf{X}_2$ est un élément de D non nul, donc inversible dans D. Il existe donc $d' \in D$ tel que

$$dd' = d'd = 1_D = 1_{\mathrm{M}_n(\mathbb{C})} = \mathrm{I}_n.$$

En particulier, $d = \mathbf{X}_1 - \mathbf{X}_2$ est alors inversible[4] dans $\mathrm{M}_n(\mathbb{C})$, donc de déterminant non nul. Puisque l'ensemble $\mathcal{C}$ est fini, l'infimum est atteint, et l'on en déduit que $\delta_{\min}(\mathcal{C}) > 0$. □

Le problème de codage Wifi revient donc à trouver en premier lieu un sous-anneau de $\mathrm{M}_n(\mathbb{C})$, qui soit aussi un anneau à division. En fait, il faut même rajouter une petite condition supplémentaire, à savoir que $k\mathrm{I}_n \subset D$, où k est un sous-corps de $\mathbb{C}$ bien choisi. Pour expliquer pourquoi, il faut tout d'abord dire que les coefficients de la matrice $\mathbf{X} \in \mathcal{C}$ que l'on envoie ne sont pas la véritable information que l'on cherche à transmettre. En fait, les symboles d'informations que l'on cherche à envoyer sont des éléments de k. Ces symboles vont alors être « encodés » dans les coefficients de la matrice $\mathbf{X}$ que l'on transmettra ensuite. Tout le jeu consiste alors à trouver une manière d'encoder de façon à obtenir un code performant (i.e. avec une probabilité d'erreur la plus petite possible).

Avant de continuer, remarquons que pour tout sous-corps k de $\mathbb{C}$, le produit matriciel induit une loi externe

$$\begin{aligned} k \times \mathrm{M}_n(\mathbb{C}) &\longrightarrow \mathrm{M}_n(\mathbb{C}) \\ (\lambda, \mathbf{X}) &\longmapsto (\lambda \mathrm{I}_n) \cdot \mathbf{X} \end{aligned}$$

qui fait de $\mathrm{M}_n(\mathbb{C})$ une k-algèbre associative unitaire[5]. Tout sous-anneau D de $\mathrm{M}_n(\mathbb{C})$ contenant $k\mathrm{I}_n$ est alors une sous-algèbre de la k-algèbre $\mathrm{M}_n(\mathbb{C})$.

Supposons maintenant qu'un tel sous-anneau D soit de dimension finie m sur k. On peut ainsi écrire tout élément de D dans une base, dont les m

4. Puisque le produit dans D est le produit matriciel, par définition d'un sous-anneau.
5. Les vérifications sont lâchement laissées au lecteur.

coordonnées dans la dite base seront des éléments de k. Cela permet alors d'encoder m symboles d'informations dans un élément de D. En pratique, pour des raisons techniques que nous n'évoquerons pas ici, on prend toujours $k = \mathbb{Q}[i]$ ou $\mathbb{Q}[j]$.

Il reste à trouver des exemples de sous-anneaux D contenus dans $\mathrm{M}_n(\mathbb{C})$ vérifiant toutes les conditions précédentes. Le but du prochain paragraphe est justement de donner toute une famille d'exemples dans le cas $n = 2$, cas auquel nous nous bornerons maintenant par souci de simplification. Le lecteur intéressé par le cas général pourra consulter [12].

2. Recherche à corps perdu

Soit k un sous-corps de $\mathbb{C}$. Dans ce paragraphe, nous allons donner une famille de sous-k-algèbres de $\mathrm{M}_2(\mathbb{C})$ qui satisfont les conditions voulues.

2.1. Théorème. *Soient $a, b \in k^\times$. Le sous-ensemble $(a,b)_k$ de $\mathrm{M}_2(\mathbb{C})$ défini par*

$$(a,b)_k = \left\{ q = \begin{pmatrix} x + z\sqrt{b} & a(y - t\sqrt{b}) \\ y + t\sqrt{b} & x - z\sqrt{b} \end{pmatrix} \mid x, y, z, t \in k \right\}$$

est une sous-k-algèbre de $\mathrm{M}_2(\mathbb{C})$ de dimension 4 non commutative. De plus, les conditions suivantes sont équivalentes :

(1) *la k-algèbre $(a,b)_k$ est un anneau à division ;*

(2) *pour tous x, y, z, $t \in k$, on a*

$$x^2 - ay^2 - bz^2 + abt^2 = 0 \implies x = y = z = t = 0$$

(3) *l'équation $a = \alpha^2 - b\beta^2$, $\alpha, \beta \in k$ n'a pas de solutions ;*

(4) *l'équation $b = \alpha^2 - a\beta^2$, $\alpha, \beta \in k$ n'a pas de solutions.*

Démonstration. On vérifie aisément que les matrices

$$\mathrm{I}_2, \quad e = \begin{pmatrix} 0 & a \\ 1 & 0 \end{pmatrix}, \quad f = \begin{pmatrix} \sqrt{b} & 0 \\ 0 & -\sqrt{b} \end{pmatrix}, \quad g = \begin{pmatrix} 0 & -a\sqrt{b} \\ \sqrt{b} & 0 \end{pmatrix}$$

forment une k-base de $(a,b)_k$, qui est donc de dimension 4. Enfin, on remarque que $g = ef$.

Un élément de $(a,b)_k$ s'écrit donc de manière unique sous la forme

$$x\mathrm{I}_n + ye + zf + tef, \quad x, y, z, t \in k.$$

Il est clair que $(a,b)_k$ contient la matrice nulle et la matrice identité, et que $(a,b)_k$ est stable par addition et par opposé. Des calculs élémentaires

montrent que l'on a

$$e^2 = a\mathrm{I}_2, \quad f^2 = b\mathrm{I}_2, \quad fe = -ef,$$

et par suite, on a aussi

$$(ef)^2 = e(fe)f = -e^2f^2 = -ab.$$

Ces relations suffisent alors à montrer que le produit de deux éléments quelconques de la base est encore un élément de $(a,b)_k$. En utilisant le fait que $(a,b)_k$ est un k-espace vectoriel et la distributivité du produit matriciel par rapport à l'addition, on en déduit que $(a,b)_k$ est stable par produit[6]. Par conséquent, $(a,b)_k$ est une sous-k-algèbre de $\mathrm{M}_2(\mathbb{C})$, non commutative puisque $ef = -fe$.

Il reste donc à voir que les conditions (1) – (4) sont équivalentes.

Montrons tout d'abord l'équivalence entre (1) et (2).

Supposons d'abord que (1) soit vérifiée, et soient $x,y,z,t \in k$ non tous nuls. Comme I_2, e, f, ef est une k-base de $(a,b)_k$, l'élément $q = x\mathrm{I}_2 + ye + zf + tef$ est non nul, donc inversible dans $(a,b)_k$ par hypothèse. Or, les égalités

$$qq^{-1} = q^{-1}q = \mathrm{I}_2 \in (a,b)_k$$

peuvent se voir comme des égalités dans $\mathrm{M}_2(\mathbb{C})$. Ainsi, q est aussi inversible dans $\mathrm{M}_2(\mathbb{C})$, et $\det(q) = x^2 - ay^2 - bz^2 + abt^2 \neq 0$, d'où (2).

Inversement, supposons que (2) soit vérifiée. Soit $q = x\mathrm{I}_2 + ye + zf + tef$ un élément non nul de $(a,b)_k$. On a donc $(x,y,z,t) \neq (0,0,0,0)$, et donc par hypothèse $x^2 - ay^2 - bz^2 + abt^2 \neq 0$, c'est-à-dire $\det(q) \neq 0$. Ainsi, la matrice q est inversible. De plus,

$$\begin{aligned} q^{-1} &= \frac{1}{x^2 - ay^2 - bz^2 + abt^2} \begin{pmatrix} x - z\sqrt{b} & -a(y - t\sqrt{b}) \\ -(y + t\sqrt{b}) & x + z\sqrt{b} \end{pmatrix} \\ &= \frac{1}{x^2 - ay^2 - bz^2 + abt^2}(x\mathrm{I}_2 - ye - zf - tef) \in (a,b)_k, \end{aligned}$$

et q est donc inversible dans $(a,b)_k$, d'où (1).

Montrons l'équivalence entre (2) et (3).

Supposons que (2) soit vérifiée. Si l'équation $a = \alpha^2 - \beta^2 b$ avait au moins une solution $(\alpha,\beta) \in k^2$, alors le quadruplet $(x,y,z,t) = (\alpha,1,\beta,0)$ serait une solution non triviale de l'équation $x^2 - ay^2 - bz^2 + abt^2 = 0$, ce qui est impossible par hypothèse, d'où (3).

Supposons que (3) soit vérifiée. Montrons tout d'abord que l'équation

$$u^2 - bv^2 = 0, \quad u,v \in k$$

6. On aurait aussi pu faire un calcul direct.

n'a pas de solution autre que $(u,v) = (0,0)$. Sinon, v est nécessairement non nul, et l'on a $b = \left(\frac{u}{v}\right)^2$. Posons $c = \frac{u}{v}$. Remarquons que c est non nul puisque b est non nul. On vérifie alors que l'on a

$$a = \left(\frac{1+a}{2}\right)^2 - b\left(\frac{1-a}{2c}\right)^2,$$

d'où une contradiction.

Soient maintenant $x, y, z, t \in k$ tels que

$$x^2 - ay^2 - bz^2 + abt^2 = 0.$$

Supposons que $(y,t) \neq (0,0)$. On a donc $y^2 - bt^2 \neq 0$, et ainsi

$$a = \frac{x^2 - bz^2}{y^2 - bt^2} = \frac{(x - bz^2)(y^2 - bt^2)}{(y^2 - bt^2)^2} = \frac{(xy + bzt)^2 - b(xt + yz)^2}{(y^2 - bt^2)^2},$$

d'où une contradiction avec (3). Ainsi, $y = t = 0$, et donc $x^2 - bz^2 = 0$, ce qui entraîne $x = z = 0$, d'où (2).

Enfin, l'équivalence entre (2) et (4) s'obtient en échangeant les rôles de a et b. Cela achève la démonstration. □

2.2. Définition. La k-algèbre $(a,b)_k$ s'appelle une *k-algèbre de quaternions* sur k.

2.3. Exemple. Grâce au théorème 2.1, il est facile de produire des exemples d'anneaux à division. Par exemple, si $k \subset \mathbb{R}$ et a et b sont strictement négatifs, alors $(a,b)_k$ est un anneau à division.

En effet, pour tous $\alpha, \beta \in k \subset \mathbb{R}$, on a $\alpha^2 - b\beta^2 \geqslant 0$, puisque $b < 0$. L'équation $a = \alpha^2 - b\beta^2$ n'a donc pas de solutions, puisque $a < 0$.

En particulier, $\mathbb{H} = (-1,-1)_{\mathbb{R}}$ est un anneau à division. Cet anneau a été découvert par Hamilton, et est le premier exemple historique d'anneau à division non commutatif.

Il est un peu moins aisé de trouver des exemples d'anneaux de quaternions qui soient à division lorsque $k = \mathbb{Q}[i]$. Le but de ce qui suit est de donner de tels exemples. On commence par un lemme.

2.4. Lemme. *Soient $u, v \in \mathbb{Z}$ tels que $p = u^2 + v^2$ soit un nombre premier, et soit $\pi = u + iv$. Alors, on a un isomorphisme d'anneaux $\mathbb{Z}[i]/(\pi) \simeq \mathbb{Z}/p\mathbb{Z}$.*

En particulier, π est premier dans $\mathbb{Z}[i]$.

Démonstration. Afin d'alléger les notations, posons $A = \mathbb{Z}[i]/(\pi)$.

Commençons par remarquer que v n'est pas divisible par p, puisque sinon, on aurait $v^2 \geqslant p^2$, puis la contradiction $p = u^2 + v^2 \geqslant p^2 > p$. Notons que

le même argument montre que u n'est pas divisible par p, même si nous n'en aurons pas besoin ici. Il existe donc $\alpha, \beta \in \mathbb{Z}$ tels que $\alpha p + \beta v = 1$. Comme $p = u^2 + v^2 = (u+iv)(u-iv)$, on a $\overline{p} = \overline{0}$. On a ainsi $\overline{\beta}\overline{v} = \overline{1} \in A$, soit encore $\overline{\beta} = \overline{v}^{-1} \in A$. Puisque $\overline{u} + \overline{i}\overline{v} = \overline{u+iv} = \overline{0} \in A$, on obtient aisément $\overline{i} = -\overline{u}\overline{\beta} = \overline{-u\beta} \in A$.

On en déduit que pour tous $r, s \in \mathbb{Z}$, on a $\overline{r+is} = \overline{r-su\beta} \in A$. Par conséquent, le morphisme d'anneaux

$$\begin{aligned} \Theta_A : \mathbb{Z} &\longrightarrow A \\ m &\longmapsto m \cdot 1_A = \overline{m} \end{aligned}$$

est surjectif. Montrons que $\mathrm{Ker}(\Theta_A) = p\mathbb{Z}$, ce qui permettra de conclure par le théorème de factorisation. Cela revient à montrer que $\mathrm{car}(A) = p$, par définition.

Notons que $p \cdot 1_A = \overline{p} = \overline{0}$, si bien que $\mathrm{car}(A) \mid p$. On a donc $\mathrm{car}(A) = 1$ ou p. L'égalité $\mathrm{car}(A) = 1$ revient à dire que l'anneau A est trivial, c'est-à-dire que $(u+iv) = A$, ou encore que $1 \in (u+iv)$. Supposons que cela soit le cas, et qu'il existe $z \in \mathbb{Z}[i]$ tel que $1 = (u+iv)z$. En comparant les modules au carré, on obtient alors la contradiction $1 = p|z|^2 \in \mathbb{Z}$. Bref, $\mathrm{car}(A) = p$, et l'on a l'isomorphisme souhaité.

Enfin, comme $\mathbb{Z}/p\mathbb{Z}$ est intègre, (π) est un idéal premier, d'où la dernière partie. □

On a alors le résultat suivant.

2.5. Proposition. *Soient $u, v \in \mathbb{Z}$ tels que $p = u^2 + v^2$ soit un nombre premier congru à 5 modulo 8. Alors, la $\mathbb{Q}[i]$-algèbre de quaternions $(i,p)_{\mathbb{Q}[i]}$ est un anneau à division.*

Démonstration. D'après le théorème 2.1, il suffit de montrer que l'équation

$$i = \alpha^2 - p\beta^2, \quad \alpha, \beta \in \mathbb{Q}[i]$$

n'a pas de solutions. Supposons le contraire. En multipliant par le carré d'un élément $z_3 \in \mathbb{Z}[i]$ bien choisi (par exemple, un dénominateur commun des parties réelles et imaginaires de α et β), on en déduit l'existence de deux éléments $z_1, z_2 \in \mathbb{Z}[i]$ vérifiant

$$iz_3^2 = z_1^2 - pz_2^2.$$

Soit $\pi = u + iv$. Si $\pi \mid z_3$ dans $\mathbb{Z}[i]$, alors $\pi \mid z_1^2$ (puisque $\pi \mid p$). Comme π est premier dans $\mathbb{Z}[i]$ d'après le lemme précédent, on a $\pi \mid z_1$. Mais alors, on a $\pi^2 \mid z_1^2$. Comme $\pi^2 \mid z_3^2$, on obtient $\pi^2 \mid pz_2^2$, soit $\pi \mid \overline{\pi} z_2^2$. Comme π

est premier dans $\mathbb{Z}[i]$, on a $\pi \mid \overline{\pi}$ ou $\pi \mid z_2^2$. Si l'on avait $\pi \mid \overline{\pi}$, on aurait

$$\frac{\overline{\pi}}{\pi} = \frac{(u-iv)^2}{p} = \frac{u^2-v^2}{p} + i\frac{2uv}{p} \in \mathbb{Z}[i].$$

En particulier, on aurait $p \mid 2uv$ dans $\mathbb{Z}$. Comme p est impair, cela impliquerait $p \mid u$ ou $p \mid v$, ce qui n'est pas le cas, comme on l'a déjà vu dans la démonstration du lemme précédent. On a donc finalement $\pi \mid z_2$.

Bref, on peut donc diviser z_1, z_2 et z_3 par π. En itérant le processus autant de fois que nécessaire, on peut alors supposer que $\pi \nmid z_3$. Dans ce cas, z_3 est non nul modulo π, donc inversible modulo π, puisque $\mathbb{Z}[i]/(\pi)$ est un corps d'après le lemme précédent. Posons $\overline{\alpha} = \overline{z}_1\overline{z}_3^{-1}$. On a ainsi $\overline{i} = \overline{\alpha}^2$, si bien que $\overline{\alpha}^4 = \overline{-1}$. Notons que $\overline{-1} \neq \overline{1} \in \mathbb{Z}[i]/(\pi)$, car $\mathbb{Z}[i]/(\pi)$ est de caractéristique p, donc impaire, toujours d'après le lemme précédent. On en déduit alors que $\overline{\alpha}$ est un élément de $(\mathbb{Z}[i]/(\pi))^\times$ d'ordre 8. Par conséquent, le groupe $(\mathbb{Z}/p\mathbb{Z})^\times$ contient aussi un élément d'ordre 8. Or, ce groupe étant cyclique d'ordre $p-1$, on en déduit que $8 \mid p-1$, soit $p \equiv 1\ [8]$, d'où une contradiction. Cela achève la démonstration. □

Revenons à notre problème de codage. Par la proposition précédente, l'algèbre de quaternions $(i,p)_{\mathbb{Q}[i]}$ est un anneau à division. En particulier, tout sous-ensemble fini $\mathcal{C} \subset (i,p)_{\mathbb{Q}[i]}$, où $p \equiv 5[8]$ est un nombre premier qui est somme de deux carrés [7], vérifie $\delta_{\min}(\mathcal{C}) > 0$.

A priori, notre problème est enfin résolu. Mais en fait... non ! Comme nous l'avons expliqué dans le premier paragraphe, on a une quantité d'énergie déterminée à l'avance disponible pour transmettre l'information. La quantité d'énergie nécessaire pour envoyer un symbole d'information $z \in \mathbb{Q}[i]$ est mesurée par $|z|^2$. Pour comprendre intuitivement pourquoi, il faut imaginer z comme un point de $\mathbb{R}^2$. La quantité d'énergie nécessaire est alors mesurée par la distance à l'origine. Plus elle est grande, plus il est coûteux d'envoyer z.

Lors d'une transmission, le coût énergétique de l'envoi de $x, y, z, t \in \mathbb{Q}[i]$ est donc

$$|x|^2 + |y|^2 + |z|^2 + |t|^2.$$

Mais ce n'est pas x, y, z, t que l'on envoie réellement, mais les coefficients de la matrice $\mathbf{X} = \begin{pmatrix} x + z\sqrt{b} & a(y - t\sqrt{b}) \\ y + t\sqrt{b} & x - z\sqrt{b} \end{pmatrix}$.

Le coût énergétique réel lors de la transmission de $\mathbf{X}$ est donc ici

$$|x + z\sqrt{b}|^2 + |a(y - t\sqrt{b})|^2 + |y + t\sqrt{b}|^2 + |x - z\sqrt{b}|^2.$$

7. En fait, cette condition est redondante. On peut en effet démontrer que tout nombre premier congru à 1 modulo 4 est une somme de deux carrés.

Pour s'assurer que l'on dispose suffisamment d'énergie, il faut donc que l'encodage préserve le coût énergétique. Autrement dit, on veut avoir

$$|x|^2+|y|^2+|z|^2+|t|^2=|x+z\sqrt{b}|^2+|a(y-t\sqrt{b})|^2+|y+t\sqrt{b}|^2+|x-z\sqrt{b}|^2,$$

pour tous $x, y, z, t \in \mathbb{Q}[i]$.

Hélas, cela est faux en général, comme on peut le voir par exemple en prenant pour b un entier strictement plus grand que 1, $x = y = t = 0$ et z arbitraire. Il faut donc modifier l'encodage de façon à satisfaire cette contrainte énergétique. C'est le but du paragraphe suivant.

3. Conservons notre énergie !

Ici, $k = \mathbb{Q}[i]$. Considérons une algèbre de quaternions $(a, b)_k$ qui soit un anneau à division. En particulier, b n'est pas un carré dans k (contempler la propriété (4) du théorème 2.1 pour s'en convaincre).

Si l'on regarde attentivement la matrice $\mathbf{X} = \begin{pmatrix} x+z\sqrt{b} & a(y-t\sqrt{b}) \\ y+t\sqrt{b} & x-z\sqrt{b} \end{pmatrix}$, on voit que ses coefficients sont des éléments du corps $k[\sqrt{b}]$. Ce corps est d'ailleurs un k-espace vectoriel de dimension 2, de base $(1, \sqrt{b})$. D'autre part, si l'on pose

$$(x_1 + x_2\sqrt{b})^* = x_1 - x_2\sqrt{b}, \quad \text{pour tous } x_1, x_2 \in k,$$

alors la matrice $\mathbf{X}$ est en fait de la forme

$$\begin{pmatrix} q_0 & aq_1^* \\ q_1 & q_0^* \end{pmatrix}, \quad q_0, q_1 \in k[\sqrt{b}].$$

Remarquons pour plus tard que l'on a les propriétés suivantes, qui se démontrent par simple calcul :

(i) pour tous $q_0, q_1 \in k[\sqrt{b}]$, on a $(q_0 + q_1)^* = q_0^* + q_1^*$;

(ii) pour tous $q_0, q_1 \in k[\sqrt{b}]$, on a $(q_0 q_1)^* = q_0^* q_1^*$;

(iii) pour tout $q \in k[\sqrt{b}]$, $(q^*)^* = q$;

(iv) pour tout $q \in k[\sqrt{b}]$, et tout $\lambda \in k$, on a $(\lambda q)^* = \lambda q^*$.

L'idée est alors la suivante : au lieu d'utiliser la base $(1, \sqrt{b})$ pour écrire les éléments de $k[\sqrt{b}]$ et encoder les symboles d'information x, y, z, t, on utilise une autre base (ω_1, ω_2), devant satisfaire des conditions encore à déterminer. On pourra toujours encoder quatre symboles d'information dans une matrice, à savoir les coordonnées de q_0 et q_1 dans cette base.

Autrement dit, on change la manière d'encoder, et l'on va plutôt envoyer

$$\mathbf{X} = \begin{pmatrix} x\omega_0 + z\omega_1 & a(y\omega_0 + t\omega_1)^* \\ y\omega_0 + t\omega_1 & (x\omega_0 + z\omega_1)^* \end{pmatrix} = \begin{pmatrix} x\omega_0 + z\omega_1 & a(y\omega_0^* + t\omega_1^*) \\ y\omega_0 + t\omega_1 & x\omega_0^* + z\omega_1^* \end{pmatrix}.$$

Remarquons que cette matrice $\mathbf{X}$ est encore un élément de $(a,b)_k$. On a juste fait un changement de base. La condition énergétique demande alors que la quantité $|x|^2 + |y|^2 + |z|^2 + |t|^2$ soit égale à

$$|x\omega_0 + z\omega_1|^2 + |a(y\omega_0^* + t\omega_1^*)|^2 + |y\omega_0 + t\omega_1|^2 + |x\omega_0^* + z\omega_1^*|^2,$$

pour tous $x, y, z, t \in k$.

Cela n'est pas très parlant. Heureusement, on peut exprimer les choses de manière plus sympathique.

3.1. Proposition. *La condition énergétique est vérifiée si, et seulement si, les deux propriétés suivantes sont vérifiées :*

(1) *on a* $|a|^2 = 1$;

(2) *la matrice* $W = \begin{pmatrix} \omega_0 & \omega_1 \\ \omega_0^* & \omega_1^* \end{pmatrix} \in \mathrm{M}_2(\mathbb{C})$ *est unitaire.*

Démonstration. La condition énergétique est vérifiée si, et seulement si, les vecteurs

$$\begin{pmatrix} x \\ z \\ y \\ t \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} x\omega_0 + z\omega_1 \\ x\omega_0^* + z\omega_1^* \\ y\omega_0 + t\omega_1 \\ a(y\omega_0^* + t\omega_1^*) \end{pmatrix}$$

ont même norme pour tous $x, y, z, t \in k$. Or, on a l'égalité

$$\begin{pmatrix} x\omega_0 + z\omega_1 \\ x\omega_0^* + z\omega_1^* \\ y\omega_0 + t\omega_1 \\ a(y\omega_0^* + t\omega_1^*) \end{pmatrix} = \begin{pmatrix} \omega_0 & \omega_1 & 0 & 0 \\ \omega_0^* & \omega_1^* & 0 & 0 \\ 0 & 0 & \omega_0 & \omega_1 \\ 0 & 0 & a\omega_0^* & a\omega_1^* \end{pmatrix} \begin{pmatrix} x \\ z \\ y \\ t \end{pmatrix},$$

pour tous $x, y, z, t \in k$.

La condition énergétique se traduit donc de la manière suivante : la matrice

$$M = \begin{pmatrix} \omega_0 & \omega_1 & 0 & 0 \\ \omega_0^* & \omega_1^* & 0 & 0 \\ 0 & 0 & \omega_0 & \omega_1 \\ 0 & 0 & a\omega_0^* & a\omega_1^* \end{pmatrix} \in \mathrm{M}_4(\mathbb{C})$$

est unitaire (i.e. ses colonnes forment une base orthonormée de $\mathbb{C}^4$ pour le produit scalaire hermitien usuel). Si l'on pose $W = \begin{pmatrix} \omega_0 & \omega_1 \\ \omega_0^* & \omega_1^* \end{pmatrix}$, on re-

marque alors que l'on a

$$M = \begin{pmatrix} W & 0 \\ 0 & \begin{pmatrix} 1 & \\ & a \end{pmatrix} W \end{pmatrix}.$$

Le résultat s'en déduit facilement. □

Le problème est que la condition (2) est rarement vérifiée, pour des raisons de déterminant. En effet, si W est unitaire, alors son déterminant est de module 1. On doit donc avoir

$$|\omega_0\omega_1^* - \omega_0^*\omega_1|^2 = |\omega_0\omega_1^* - (\omega_0\omega_1^*)^*|^2 = 1.$$

Imaginons que b soit un entier positif. Si l'on écrit

$$\omega_0\omega_1^* = \alpha + \beta\sqrt{b}, \quad \alpha, \beta \in k,$$

la condition précédente se traduit par $4|\beta|^2 b = 1$. Si $\beta = r + si$, $r, s \in \mathbb{Q}$, on obtient alors

$$b = \frac{r^2 + s^2}{4(r^2 + s^2)^2}.$$

Ainsi, b doit être une somme de deux carrés de $\mathbb{Q}$, ce qui n'est pas forcément le cas. Pour franchir cet obstacle, on modifie encore un peu l'encodage.

On choisit un élément $\lambda \in k[\sqrt{b}] \cap \mathbb{R}$ tel que $\lambda > 0$ et $\lambda^* > 0$. On envoie alors, non pas la matrice $\mathbf{X} = \begin{pmatrix} x\omega_0 + z\omega_1 & a(y\omega_0 + t\omega_1)^* \\ y\omega_0 + t\omega_1 & (x\omega_0 + z\omega_1)^* \end{pmatrix}$, mais la matrice

$$\mathbf{X}_\lambda = \begin{pmatrix} \sqrt{\lambda} & 0 \\ 0 & \sqrt{\lambda^*} \end{pmatrix} \mathbf{X} = \begin{pmatrix} \sqrt{\lambda}(x\omega_0 + z\omega_1) & a\sqrt{\lambda^*}(y\omega_0 + t\omega_1)^* \\ \sqrt{\lambda}(y\omega_0 + t\omega_1) & \sqrt{\lambda^*}(x\omega_0 + z\omega_1)^* \end{pmatrix}.$$

Des calculs semblables à ceux effectués précédemment montrent que la condition énergétique est vérifiée si, et seulement si, les deux propriétés suivantes sont vérifiées :

(1) on a $|a|^2 = 1$;
(2) la matrice $W_\lambda = \begin{pmatrix} \sqrt{\lambda}\omega_0 & \sqrt{\lambda}\omega_1 \\ \sqrt{\lambda^*}\omega_0^* & \sqrt{\lambda^*}\omega_1^* \end{pmatrix} \in \mathrm{M}_2(\mathbb{C})$ est unitaire.

Pour des raisons que nous n'exposerons pas ici, cette dernière modification permet d'augmenter les chances pour que la condition (2) soit vérifiée. Remarquer aussi que par définition, les matrices $\mathbf{X}_\lambda$ sont de la forme

$$\begin{pmatrix} \sqrt{\lambda} & 0 \\ 0 & \sqrt{\lambda^*} \end{pmatrix} \mathbf{X}, \quad \text{avec } \mathbf{X} \in (a, b)_k.$$

Ainsi, la différence de deux matrices $\mathbf{X}_\lambda, \mathbf{X}'_\lambda$ distinctes est inversible, dès que $(a, b)_k$ est un anneau à division.

En résumé, on a montré la chose suivante : si les deux conditions précédentes sont vérifiées, alors tout sous-ensemble $\mathcal{C}$ de l'ensemble

$$\mathcal{C}_\lambda = \left\{ \mathbf{X}_\lambda = \begin{pmatrix} \sqrt{\lambda}(x\omega_0 + z\omega_1) & a\sqrt{\lambda^*}(y\omega_0 + t\omega_1)^* \\ \sqrt{\lambda}(y\omega_0 + t\omega_1) & \sqrt{\lambda^*}(x\omega_0 + z\omega_1)^* \end{pmatrix} \;\middle|\; x, y, z, t \in k \right\}$$

vérifie $\delta_{\min}(\mathcal{C}) > 0$, et de plus, l'encodage préserve le coût énergétique.

Donnons maintenant une famille d'exemples.

3.2. Lemme. *Soit $k = \mathbb{Q}[i]$, et soient $u, v \in \mathbb{Z}$ tels que $p = u^2 + v^2$ est un nombre premier congru à 5 modulo 8. Soient*

$$\lambda = \frac{1}{p}, \quad \omega_0 = \frac{u + iv + \sqrt{p}}{2}, \quad \omega_1 = \frac{u + iv - \sqrt{p}}{2}.$$

Alors, (ω_0, ω_1) est une k-base de $k[\sqrt{p}]$, et la matrice W_λ est unitaire.

Démonstration. C'est un simple calcul. □

En explicitant les matrices, on a alors le résultat suivant.

3.3. Corollaire. *Soit $k = \mathbb{Q}[i]$. Soient $u, v \in \mathbb{Z}$ tels que $p = u^2 + v^2$ est un nombre premier congru à 5 modulo 8, et soit $\mathcal{C}_{\frac{1}{p}}$ l'ensemble des matrices de la forme*

$$\frac{1}{2\sqrt{p}} \begin{pmatrix} (x+z)(u+iv) + (x-z)\sqrt{p} & i\big((y+t)(u+iv) - (y-t)\sqrt{p}\big) \\ (y+t)(u+iv) + (y-t)\sqrt{p} & (x+z)(u+iv) - (x-z)\sqrt{p} \end{pmatrix},$$

avec $x, y, z, t \in k$.

Alors, pour tout sous-ensemble fini $\mathcal{C}$ de $\mathcal{C}_{\frac{1}{p}}$, on a $\delta_{\min}(\mathcal{C}) > 0$. De plus, l'encodage des symboles d'informations $x, y, z, t \in k$ préserve le coût énergétique.

A-t-on fini, cette fois ? Non, toujours pas. Idéalement, on aimerait que les performances du code $\mathcal{C}$ ne dépendent pas de la taille de son cardinal. Autrement dit, on aimerait que le déterminant minimum de $\mathcal{C}$ soit suffisamment grand, que l'on prenne 16, 32 ou 2^{64} matrices. Or, si $\mathcal{C}$ est mal choisi, il pourrait contenir des matrices $\mathbf{X}_1$ et $\mathbf{X}_2$ telle que $|\det(\mathbf{X}_1 - \mathbf{X}_2)|^2$ soit très proche de 0 (par exemple, on pourrait avoir $\mathbf{X}_2 = \mathbf{X}_1 + \frac{1}{n} I_2$), ce qui donnerait un déterminant minimum très proche de 0 également. Pour pallier cet inconvénient, une solution consiste à forcer $|\det(\mathbf{X}_1 - \mathbf{X}_2)|^2$ à prendre ses valeurs dans $\mathbb{Z}$ (ou dans un ensemble de la forme $c\mathbb{Z}, c \in \mathbb{R}$). L'idée est de prendre $x, y, z, t \in \mathbb{Z}[i]$.

4. Soyons discrets...

On reprend les notations du paragraphe précédent. Posons alors

$$\mathcal{C}'_\lambda = \left\{ \mathbf{X}_\lambda = \begin{pmatrix} \sqrt{\lambda}(x\omega_0 + z\omega_1) & a\sqrt{\lambda^*}(y\omega_0 + t\omega_1)^* \\ \sqrt{\lambda}(y\omega_0 + t\omega_1) & \sqrt{\lambda^*}(x\omega_0 + z\omega_1)^* \end{pmatrix} \;\middle|\; x, y, z, t \in \mathbb{Z}[i] \right\}.$$

On a alors le résultat suivant.

4.1. Proposition. *Il existe un réel $\alpha > 0$ tel que pour tout sous-ensemble fini $\mathcal{C} \subset \mathcal{C}'_\lambda$, on a $\delta_{\min}(\mathcal{C}) \geqslant \alpha$.*

Démonstration. Choisissons un entier $d \geqslant 1$ tel que $d\omega_0, d\omega_1 \in \mathbb{Z}[i][\sqrt{b}]$ (un tel entier existe toujours). Alors, pour tout $\mathbf{X}_\lambda \in \mathcal{C}'_\lambda$, on a

$$d\mathbf{X}_\lambda = \begin{pmatrix} \sqrt{\lambda}(z_0 + z_2\sqrt{b}) & a\sqrt{\lambda^*}(z_1 - z_3\sqrt{b}) \\ \sqrt{\lambda}(z_1 + z_3\sqrt{b}) & \sqrt{\lambda^*}(z_0 - z_2\sqrt{b}) \end{pmatrix},$$

pour certains $z_0, z_1, z_2, z_3 \in \mathbb{Z}[i]$. Choisissons maintenant un entier $c \geqslant 1$ tel que $ca, cb \in \mathbb{Z}[i]$. On a alors

$$\det(cd\mathbf{X}_\lambda) = c^2 \det(d\mathbf{X}_\lambda) = \sqrt{\lambda\lambda^*}c^2(z_0^2 - az_1^2 - bz_2^2 + abz_3^2) \in \sqrt{\lambda\lambda^*}\mathbb{Z}[i].$$

En particulier, $|\det(cd\mathbf{X}_\lambda)|^2 \in \lambda\lambda^*\mathbb{N}$. Ainsi, on a $|\det(cd\mathbf{X}_\lambda)|^2 \geqslant \lambda\lambda^*$ dès que $\mathbf{X}_\lambda \neq 0$ (les éléments non nuls de $\mathcal{C}_\lambda$, et donc de $\mathcal{C}'_\lambda$, sont inversibles par construction). On a donc

$$|\det(\mathbf{X}_\lambda)|^2 \geqslant \frac{\lambda\lambda^*}{(cd)^4}, \quad \text{pour tout } \mathbf{X}_\lambda \in \mathcal{C}'_\lambda \text{ non nul.}$$

Remarquons maintenant que $\mathcal{C}'_\lambda$ est stable par addition et par opposé. Pour tous $\mathbf{X}_\lambda, \mathbf{X}'_\lambda \in \mathcal{C}'_\lambda$ distincts, on a alors

$$|\det(\mathbf{X}_\lambda - \mathbf{X}'_\lambda)|^2 \geqslant \frac{\lambda\lambda^*}{(cd)^4}.$$

On en déduit que $\delta_{\min}(\mathcal{C}'_\lambda) \geqslant \dfrac{\lambda\lambda^*}{(cd)^4}$. Par suite, pour tout $\mathcal{C} \subset \mathcal{C}'_\lambda$, on obtient

$$\delta_{\min}(\mathcal{C}) \;\geqslant\; \delta_{\min}(\mathcal{C}'_\lambda) \;\geqslant\; \frac{\lambda\lambda^*}{(cd)^4}.$$

Cela achève la démonstration. □

Ce résultat est particulièrement intéressant, puisqu'il donne l'existence d'une borne inférieure pour le déterminant minimum de **n'importe quel** sous-ensemble fini $\mathcal{C}$ de $\mathcal{C}'_\lambda$. En particulier, cela garantit les performances du code $\mathcal{C}$, quelque soit son nombre d'éléments.

L'exemple développé dans ce chapitre montre que l'on peut prendre pour $\mathcal{C}$ un sous-ensemble de l'ensemble $\mathcal{C}'_{\frac{1}{p}}$ des matrices de la forme

$$\frac{1}{2\sqrt{p}}\begin{pmatrix}(x+z)(u+iv)+(x-z)\sqrt{p} & i\big((y+t)(u+iv)-(y-t)\sqrt{p}\big)\\ (y+t)(u+iv)+(y-t)\sqrt{p} & (x+z)(u+iv)-(x-z)\sqrt{p}\end{pmatrix},$$

où $x,y,z,t\in\mathbb{Z}[i]$ et $p=u^2+v^2$ est un nombre premier congru à 5 modulo 8. On peut d'ailleurs estimer $\delta_{\min}(\mathcal{C}'_{\frac{1}{p}})$. Puisque $p\equiv 5$ [8], on peut supposer que u est impair, et que v est pair. Si $x,y,z,t\in\mathbb{Z}[i]$, on peut montrer alors que le déterminant de la matrice

$$\frac{1}{2}\begin{pmatrix}(x+z)(u+iv)+(x-z)\sqrt{p} & i\big((y+t)(u+iv)-(y-t)\sqrt{p}\big)\\ (y+t)(u+iv)+(y-t)\sqrt{p} & (x+z)(u+iv)-(x-z)\sqrt{p}\end{pmatrix}$$

est de la forme $(u+iv)z'$, $z'\in\mathbb{Z}[i]$.

Ainsi, toute matrice $\mathbf{X}'\in\mathcal{C}'_{\frac{1}{p}}$ non nulle vérifie

$$\big|\det(\mathbf{X}')\big|^2=\big|\tfrac{1}{p}(u+iv)z'\big|^2\geqslant\frac{1}{p}.$$

En raisonnant comme dans la démonstration précédente, on obtient

$$\delta_{\min}(\mathcal{C}'_{\frac{1}{p}})\geqslant\frac{1}{p}.$$

On peut en fait démontrer que $\delta_{\min}(\mathcal{C}'_{\frac{1}{p}})=\frac{1}{p}$, mais la démonstration utilise des outils extrêmement sophistiqués. Nous allons nous borner à le montrer pour $p=5$. Dans ce cas, on a $u=1$ et $v=2$, et donc un élément de $\mathcal{C}'_{\frac{1}{5}}$ est de la forme

$$\frac{1}{2\sqrt{5}}\begin{pmatrix}(x+z)(1+2i)+(x-z)\sqrt{5} & i\big((y+t)(1+2i)-(y-t)\sqrt{5}\big)\\ (y+t)(1+2i)+(y-t)\sqrt{5} & (x+z)(1+2i)-(x-z)\sqrt{5}\end{pmatrix},$$

avec $x,y,z,t\in\mathbb{Z}[i]$. Mais on a

$$\Big|\det\frac{1}{2\sqrt{5}}\begin{pmatrix}(1+2i)+\sqrt{5} & 0\\ 0 & (1+2i)-\sqrt{5}\end{pmatrix}\Big|^2=\frac{1}{5},$$

ce qui démontre l'égalité voulue.

Le code $\mathcal{C}'_{\frac{1}{5}}$ s'appelle le *Golden code*, et a été implémenté par plusieurs opérateurs téléphoniques.

Chapitre XX

Matrices à coefficients dans un anneau euclidien

On s'intéresse dans ce chapitre aux matrices à coefficients dans un anneau euclidien, et plus précisément à l'équivalence de ces matrices. On donnera alors des applications à la résolution des systèmes linéaires à coefficients dans un anneau euclidien, ainsi qu'à la structure des groupes abéliens de type fini.

On commence par donner quelques compléments sur les matrices à coefficients dans un anneau commutatif.

Dans la suite, sauf mention expresse du contraire, A désignera un anneau commutatif.

1. Déterminant : le retour

Dans ce court paragraphe, nous allons définir le déterminant d'une matrice carrée $M \in \mathrm{M}_n(A)$, et donner ses principales propriétés.

1.1. Définition. Soit $n \geqslant 1$, et soit $M \in \mathrm{M}_n(A)$. Le *déterminant* de M est l'élément $\det(M) \in A$ défini par

$$\det(M) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}.$$

Le déterminant jouit peu ou prou des mêmes propriétés que le déterminant d'une matrice à coefficients dans un corps. Par exemple, le contenu de la proposition IV-4.2 reste vrai dans ce contexte. Pour le voir, on peut passer

par la théorie des formes n-linéaires alternées comme dans le chapitre IV, en remplaçant le K-espace vectoriel E par A^n, mais on peut aussi utiliser l'astuce qui consiste à se réduire au cas des matrices à coefficients dans un corps, comme nous allons maintenant l'expliquer.

Montrons par exemple que pour tous M, $N \in \mathrm{M}_n(A)$, on a

$$\det(MN) = \det(NM) = \det(M)\det(N).$$

Écrivons $M = (a_{ij})$ et $N = (b_{ij})$, avec a_{ij}, $b_{ij} \in A$, et considérons l'anneau de polynômes $A_0 = \mathbb{Z}[X_{ij}, Y_{ij},\ i, j \in [\![1, n]\!]]$. Puisque A est commutatif, une application répétée de la proposition XV-6.16 montre qu'il existe un unique morphisme d'anneaux $\varphi : A_0 \longrightarrow A$ tel que

$$\varphi(a)=a,\ \varphi(X_{ij})=a_{ij},\ \varphi(Y_{ij})=b_{ij}, \quad \text{pour tout } a\in A \text{ et tous } i,j\in[\![1,n]\!].$$

On vérifie alors aisément que l'application

$$\begin{aligned} \hat{\varphi} : \mathrm{M}_n(A_0) &\longrightarrow \mathrm{M}_n(A) \\ (c_{ij}) &\longmapsto \big(\varphi(c_{ij})\big) \end{aligned}$$

est un morphisme d'anneaux. Posons $M_0 = (X_{ij})$ et $N_0 = (Y_{ij})$. La définition de $\hat{\varphi}$ et le fait que φ soit un morphisme d'anneaux montrent alors que l'on a

(i) $\hat{\varphi}(M_0) = M$ et $\hat{\varphi}(N_0) = N$;

(ii) pour tout $C \in \mathrm{M}_n(A_0)$ et tout $a \in A$, on a $\hat{\varphi}(aC) = a\hat{\varphi}(C)$;

(iii) pour tout $C \in \mathrm{M}_n(A_0)$, on a $\det\big(\hat{\varphi}(C)\big) = \varphi\big(\det(C)\big)$.

Remarquons maintenant que A_0 est un anneau intègre, qui s'identifie canoniquement à un sous-anneau de son corps des fractions, que l'on notera K_0. On a donc

$$\det(M_0N_0) = \det(N_0M_0) = \det(M_0)\det(N_0),$$

où M_0 et N_0 sont vues ici comme des matrices à coefficients dans K_0. Mais, ces égalités sont en fait des égalités dans A_0, puisque le déterminant d'une matrice $C \in \mathrm{M}_n(A_0)$ et le déterminant de la même matrice vue comme un élément de $\mathrm{M}_n(K_0)$ coïncident [1].

En appliquant φ, et en utilisant que φ est un morphisme d'anneaux, on obtient

$$\varphi\big(\det(M_0N_0)\big) = \varphi\big(\det(N_0M_0)\big) = \varphi\big(\det(M_0)\big)\varphi\big(\det(N_0)\big),$$

soit encore

$$\det\big(\hat{\varphi}(M_0N_0)\big) = \det\big(\hat{\varphi}(N_0M_0)\big) = \det\big(\hat{\varphi}(M_0)\big)\det\big(\hat{\varphi}(M_0)\big)$$

par (iii).

1. Ils sont définis par la même formule.

Comme $\hat{\varphi}$ est un morphisme d'anneaux, on en déduit

$$\det\big(\hat{\varphi}(M_0)\hat{\varphi}(N_0)\big) = \det\big(\hat{\varphi}(N_0)\hat{\varphi}(M_0)\big) = \det\big(\hat{\varphi}(M_0)\big)\det\big(\hat{\varphi}(N_0)\big),$$

d'où les égalités

$$\det(MN) = \det(NM) = \det(M)\det(N)$$

d'après (i).

Les autres propriétés listées dans la proposition IV-4.2 se démontrent de manière similaire.

Si $M \in \mathrm{M}_n(A)$, pour tous $i, j \in [\![1, n]\!]$, on définit le mineur $\Delta_{ij}(M)$ comme étant le déterminant de la matrice extraite de M, obtenue en enlevant la ligne i et la colonne j.

On définit alors la comatrice de M comme étant la matrice

$$\mathrm{com}(M) = \big((-1)^{i+j}\Delta_{ij}(M)\big) \in \mathrm{M}_n(A).$$

Pour tous entiers $i, j \in [\![1, n]\!]$, et toute matrice $C \in \mathrm{M}_n(A_0)$, on a alors (avec les notations précédentes)

$$\Delta_{ij}\big(\hat{\varphi}(C)\big) = \varphi\big(\Delta_{ij}(C)\big).$$

En particulier, on a

$$\mathrm{com}\big(\hat{\varphi}(C)\big) = \hat{\varphi}\big(\mathrm{com}(C)\big).$$

En utilisant cette égalité et le fait évident que $\hat{\varphi}$ commute avec l'application transposée, on démontre alors que

$$\mathrm{com}(M)^t M = M\mathrm{com}(M)^t = \det(M)\mathrm{I}_n, \quad \text{pour tout } M \in \mathrm{M}_n(A).$$

On a alors le fait suivant, dont la démonstration est laissée en exercice au lecteur : une matrice $M \in \mathrm{M}_n(A)$ est inversible dans l'anneau $\mathrm{M}_n(A)$ si, et seulement si, $\det(M) \in A^\times$. Dans ce cas, on a

$$M^{-1} = \det(M)^{-1}\mathrm{com}(M)^t.$$

On peut alors définir les groupes $\mathrm{GL}_n(A)$ et $\mathrm{SL}_n(A)$ pour tout $n \geqslant 1$ de manière habituelle, à savoir

$$\mathrm{GL}_n(A) = \mathrm{M}_n(A)^\times = \{M \in \mathrm{M}_n(A) \mid \det(M) \in A^\times\}$$

et

$$\mathrm{SL}_n(A) = \{M \in \mathrm{M}_n(A) \mid \det(M) = 1\}.$$

Bien entendu, $\mathrm{GL}_n(A)$ est un groupe pour la multiplication matricielle, et l'ensemble $\mathrm{SL}_n(A)$ est un sous-groupe distingué de $\mathrm{GL}_n(A)$.

1.2. Remarque. Comme dans le cas des matrices à coefficients dans un corps, on peut étendre la notion de matrice inversible aux matrices rectangulaires à coefficients dans un anneau non nécessairement commutatif de la manière suivante.

Si A est un anneau quelconque, une matrice $M \in \mathrm{M}_{m\times n}(A)$ est dite inversible s'il existe $N \in \mathrm{M}_{n\times m}(A)$ telle que

$$MN = \mathrm{I}_m \quad \text{et} \quad NM = \mathrm{I}_n.$$

Si $n = m$, une matrice $M \in \mathrm{M}_n(A)$ sera donc inversible si, et seulement si, elle est inversible dans l'anneau $\mathrm{M}_n(A)$.

Lorsque A est commutatif, on démontre qu'une matrice $M \in \mathrm{M}_{m\times n}(A)$ qui est inversible est nécessairement carrée[2]. En revanche, on peut exhiber un exemple d'anneau A non commutatif et d'une matrice inversible à coefficients dans A non carrée (cf. [9, Remarque II-2.16]).

2. Équivalence de matrices

Nous commençons par définir la notion centrale de ce paragraphe.

2.1. Définition. Soient $m, n \geqslant 1$ deux entiers.

Deux matrices C, $C' \in \mathrm{M}_{m\times n}(A)$ sont dites *équivalentes* s'il existe deux matrices $U \in \mathrm{GL}_m(A)$ et $V \in \mathrm{GL}_n(A)$ telles que $C' = UCV$.

On le note $C \sim C'$.

2.2. Remarques

(1) La relation $\sim$ est clairement une relation d'équivalence.

(2) Lorsque A est un corps, on retrouve la notion usuelle.

Lorsque A est un corps K, on sait que toute matrice est équivalente à une matrice par blocs de la forme

$$\begin{pmatrix} \mathrm{I}_r & \mathrm{O}_{r,n-r} \\ \mathrm{O}_{m-r,r} & \mathrm{O}_{m-r,n-r} \end{pmatrix}.$$

Si A est un anneau commutatif quelconque, on ne peut pas espérer un résultat similaire. Par exemple, si $m = n = 1$, les matrices (a) et (1) ne sont équivalentes que lorsque $a \in A^\times$ (cela provient aisément de la définition).

Nous avons donc besoin d'introduire une famille de matrices qui joueront le rôle des matrices

$$\begin{pmatrix} \mathrm{I}_r & \mathrm{O}_{r,n-r} \\ \mathrm{O}_{m-r,r} & \mathrm{O}_{m-r,n-r} \end{pmatrix}.$$

Soit $r \geqslant 1$ un entier, et soient $a_1, \ldots, a_r \in A$ des éléments non nuls vérifiant

$$a_1 \mid a_2 \mid \cdots \mid a_r.$$

2. C'est par exemple une conséquence du corollaire II-2.21 et du lemme II-2.15 de [9].

On pose

$$E(a_1,\dots,a_r)=\begin{pmatrix} a_1 & & & \\ & \ddots & & \mathrm{O}_{r,n-r} \\ & & a_r & \\ \mathrm{O}_{m-r,r} & & & \mathrm{O}_{m-r,n-r} \end{pmatrix} \in \mathrm{M}_{m\times n}(A).$$

On convient que cette matrice est nulle si $r=0$.

Nous allons voir que, lorsque A est euclidien, toute matrice est équivalente à une matrice $E(a_1,\dots,a_r)$. Comme dans le cas classique des matrices à coefficients dans un corps, nous procéderons par opérations élémentaires sur les lignes et les colonnes. À cette fin, nous allons introduire des matrices qui nous serviront par la suite.

Soit $\ell\geqslant 2$ un entier. Si r, $s\in[\![1,\ell]\!]$, on note E_{rs} la matrice de $\mathrm{M}_\ell(A)$ qui admet un seul coefficient non nul égal 1 à l'intersection de la ligne r et de la colonne s.

Si i, $j\in[\![1,\ell]\!]$, avec $i\neq j$, on pose

$$S_{i,j}^{(\ell)}=E_{ij}+E_{ji}+\sum_{k\neq i,j}E_{kk}\in\mathrm{M}_\ell(A).$$

Autrement dit, $S_{i,j}^{(\ell)}$ est la matrice

$$\begin{array}{c} \\ i\to \\ \\ \\ j\to \\ \\ \end{array}\begin{array}{c} \begin{array}{ccccccccc} & & & i & & & j & & \\ & & & \downarrow & & & \downarrow & & \end{array} \\ \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & 0 & & & 1 & & \\ & & & & 1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & 1 & & \\ & & & 1 & & & 0 & & \\ & & & & & & & 1 & \\ & & & & & & & & \ddots \\ & & & & & & & & & 1 \end{pmatrix} \end{array} \in \mathrm{M}_\ell(A).$$

Cette matrice est un élément de $\mathrm{GL}_\ell(A)$, d'inverse

$$(S_{i,j}^{(\ell)})^{-1}=S_{i,j}^{(\ell)}.$$

Si $C\in\mathrm{M}_{m\times n}(A)$, l'opération

$$C\longmapsto S_{i,j}^{(m)}C$$

revient à échanger les lignes i et j de C, tandis que l'opération

$$C\longmapsto CS_{i,j}^{(n)}$$

revient à échanger les colonnes i et j de C. Une matrice de ce type est appelée une *matrice de transposition.*

Si $i, j \in [\![1, \ell]\!]$, avec $i \neq j$, et $a \in A$, on pose

$$T_{i,j}^{(\ell)}(a) = \mathrm{I}_\ell + aE_{ij} \in \mathrm{M}_\ell(A).$$

Autrement dit, $T_{i,j}^{(\ell)}(a)$ est la matrice

$$\begin{array}{c} \\ i \to \end{array} \begin{array}{c} \begin{array}{c} \\ j \\ \downarrow \end{array} \\ \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & a & \\ & & & \ddots \\ & & & & 1 \end{pmatrix} \end{array} \in \mathrm{M}_\ell(A).$$

Cette matrice est un élément de $\mathrm{GL}_\ell(A)$, d'inverse

$$\left(T_{i,j}^{(\ell)}(a)\right)^{-1} = T_{i,j}^{(\ell)}(-a).$$

Si $C \in \mathrm{M}_{m\times n}(A)$, l'opération

$$C \longmapsto T_{i,j}^{(m)}(a)C$$

revient à faire l'opération

$$L_i \longleftarrow L_i + aL_j,$$

où $L_1, \ldots, L_m$ sont les lignes de C, tandis que l'opération

$$C \longmapsto CT_{i,j}^{(n)}(a)$$

revient à faire l'opération

$$C_j \longleftarrow C_j + aC_i,$$

où $C_1, \ldots, C_n$ sont les colonnes de C. Une matrice de ce type est appelée une *matrice de transvection.*

Enfin, si $u \in A^\times$, on note $D^{(\ell)}(u)$ la matrice

$$D^{(\ell)}(u) = \begin{pmatrix} \mathrm{I}_{\ell-1} & \\ & u \end{pmatrix} \in \mathrm{M}_\ell(A).$$

Cette matrice est un élément de $\mathrm{GL}_\ell(A)$, d'inverse

$$\left(D^{(\ell)}(u)\right)^{-1} = D^{(\ell)}(u^{-1}).$$

Une matrice de ce type est appelée une *matrice de dilatation.*

Si $C \in \mathrm{M}_{m\times n}(A)$, l'opération

$$C \longmapsto D^{(m)}(u)C$$

revient à multiplier la dernière ligne de C par u, tandis que l'opération

$$C \longmapsto CD^{(n}(u)$$

revient à multiplier la dernière colonne de C par u.

On a alors le théorème suivant.

2.3. Théorème (Réduction des matrices). *Soit (A, δ) un anneau euclidien, et soit $\mathscr{P}$ un s.c.r.i. de A. Toute matrice $C \in \mathrm{M}_{m\times n}(A)$ est A-équivalente à une matrice de la forme $E(a_1, \ldots, a_r)$, avec $a_1, \ldots, a_r \in A$ non nuls et normalisés vérifiant*

$$a_1 \mid a_2 \mid \cdots \mid a_r.$$

De plus, l'entier r et les éléments $a_1, \ldots, a_r$ sont uniques.

Plus précisément, il existe des matrices $U \in \mathrm{GL}_m(A)$ et $V \in \mathrm{GL}_n(A)$ qui sont produits de matrices de permutation, de transvection et de dilatation telles que $UCV = E(a_1, \ldots, a_r)$.

De plus, r est le rang de la matrice C (vue comme élément de $\mathrm{M}_{m\times n}(K_A)$, où K_A est le corps des fractions de A), et pour tout $j \in [\![0, r]\!]$, on a

$$\mu_j(C) = a_1 \cdots a_j,$$

où $\mu_j(C)$ désigne le pgcd normalisé des mineurs d'ordre j de C (avec la convention $\mu_0(C) = 1$).

En particulier, on a $a_j = \dfrac{\mu_j(C)}{\mu_{j-1}(C)}$, pour tout $j \in [\![1, r]\!]$.

Démonstration. Soit $C \in \mathrm{M}_{m\times n}(A)$. On commence par démontrer l'existence d'une matrice $E(a_1, \ldots, a_r)$ équivalente à C. Si $C = 0$, il n'y a rien à faire. On suppose donc que C est non nulle.

Nous allons démontrer de manière algorithmique que C est équivalente à une matrice de la forme

$$\begin{pmatrix} a & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & C' & \\ 0 & & & \end{pmatrix},$$

où $a \in A$ est normalisé et divise tous les coefficients de C', ce qui permettra de conclure par récurrence.

Algorithme de réduction des matrices

Étape 1. On place un coefficient non nul b tel que $\delta(b)$ soit minimal en haut à gauche, en utilisant des échanges de lignes et de colonnes (ce qui revient à multiplier à gauche ou à droite par des matrices de transposition.

Étape 2. Si b divise tous les éléments de la première ligne et de la première colonne, on passe à l'étape suivante. Sinon, on repère un élément a de la première ligne ou de la première colonne non divisible par b. En faisant des opérations du type $L_k \longleftarrow L_k + \alpha L_1$ ou du type $C_k \longleftarrow C_k + \alpha C_1$ (avec $\alpha \in A$ et $k \geqslant 2$), après échange éventuels de lignes ou de colonnes, on remplace b par le reste de la division euclidienne de a par b, disons b_1. Quitte échanger L_1 et L_m, multiplier par une matrice de dilatation, et échanger à nouveau L_1 et L_m, on peut supposer b_1 normalisé. Remarquons que $\delta(b_1) < \delta(b)$ (car b ne divise pas a).

Si tous les éléments de la première ligne et de la première colonne sont divisibles par b_1, on passe à l'étape suivante. Sinon, on recommence le procédé pour remplacer b_1 par un élément b_2 tel que $\delta(b_2) < \delta(b_1)$. Si b_2 convient, on s'arrête, sinon on recommence. Notons que le procédé s'arrête nécessairement puisque la suite $\delta(b)$, $\delta(b_1)$, $\delta(b_2), \ldots$ est strictement décroissante.

Étape 3. On utilise des opérations sur les lignes et les colonnes afin de faire apparaître des zéros et obtenir une matrice de la forme

$$\begin{pmatrix} d_1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & C_1 & \\ 0 & & & \end{pmatrix},$$

avec d_1 non nul et normalisé. Si d_1 divise tous les éléments de C_1, on a fini. Sinon, on repère une ligne i de C_1 contenant un élément non divisible par d_1, et l'on fait l'opération

$$L_1 \longleftarrow L_1 + L_i.$$

On repart alors au début de l'algorithme pour obtenir une matrice de la forme

$$\begin{pmatrix} d_2 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & C_2 & \\ 0 & & & \end{pmatrix},$$

avec d_2 non nul, normalisé et $\delta(d_2) < \delta(d_1)$. Si d_2 divise touts les coefficients de C_2, on s'arrête, sinon on recommence en travaillant avec la matrice C_2. Pour les mêmes raisons qu'avant, le procédé s'arrête nécessairement, et l'on finit par obtenir une matrice de la forme voulue.

Les opérations élémentaires sur les lignes et les colonnes se traduisant par la multiplication à gauche et à droite par des matrices de permutation, transvection ou dilatation, cela démontre l'existence de U et V.

Montrons la dernière partie. Il est clair que r est le rang de C. Avant de poursuivre, nous allons montrer le pgcd normalisé de tous les mineurs

d'ordre k de C est invariant par opérations élémentaires sur les lignes et les colonnes, et cela pour tout $k \in [\![1, \min(m,n)]\!]$.

Commençons par examiner l'effet d'une opération du type $L_i \longleftarrow L_i + aL_j$. Soit M une matrice extraite de C de taille k. Si M ne contient ni L_i, ni L_j, ou contient uniquement L_j, cette opération laisse cette matrice extraite inchangée. Si M contient L_i et L_j, cette opération induit une opération du même type, ce qui ne change pas son déterminant. Enfin, si M ne contient que la ligne L_i, le déterminant de la nouvelle matrice est de la forme $\det(M) + a\det(M')$, où M' est une matrice de taille k qui est obtenue par permutation des lignes d'une autre matrice extraite. Il en est de même pour les opérations du type $C_j \longleftarrow C_j + aC_i$. Multiplier par une matrice de dilatation soit ne change pas un mineur, soit le multiplie par un élément de $A^\times$. Regardons enfin ce qui se passe lorsque l'on échange deux lignes, le cas de l'échange de deux colonnes se démontrant de la même manière. Considérons une matrice extraite de taille k. Si M contient les deux lignes échangées ou si M ne contient aucune des deux lignes échangées, on obtient le même mineur, éventuellement au signe près. Si M contient exactement une des deux lignes échangées, elle est transformée en une matrice de taille k qui est obtenue par permutation des lignes d'une autre matrice extraite M'. En particulier, une telle opération envoie un mineur de taille k sur un autre mineur, au signe près. Dans tous les cas, le pgcd des mineurs de taille k reste alors inchangé.

Pour tout $j \in [\![0, r]\!]$, on a donc $\mu_j(C) = \mu_j\big(E(a_1, \ldots, a_r)\big)$. On voit facilement que $\mu_j\big(E(a_1, \ldots, a_r)\big)$ est le pgcd unitaire des polynômes

$$a_{i_1} \cdots a_{i_j}, \quad 1 \leqslant i_1 < \cdots < i_j \leqslant r.$$

Les relations de divisibilité $a_1 \mid a_2 \mid \cdots \mid a_r$ impliquent alors que ce pgcd est égal à $a_1 \cdots a_j$. En particulier, on a

$$\frac{\mu_j(C)}{\mu_{j-1}(C)} = \frac{a_1 \cdots a_j}{a_1 \cdots a_{j-1}} = a_j.$$

On en déduit la dernière partie du théorème. □

2.4. Remarques

(1) Si $C \in \mathrm{M}_{m\times n}(A)$, le théorème 2.3 nous permet de calculer des éléments $a_1, \ldots, a_r \in A$ tels que $C \sim E(a_1, \ldots, a_r)$ sans utiliser l'algorithme. Néanmoins, cet algorithme a son importance, car il permet de calculer des matrices U et V de manière explicite, qu'il est parfois utile de connaître, comme on le verra plus loin.

(2) Les matrices de transposition ne sont utilisées que pour ramener rapidement un élément non nul en bas à droite. Elles sont néanmoins inutiles, car l'on peut échanger deux lignes ou deux colonnes au signe près

en utilisant seulement des matrices de transvection. Plus précisément, si $\varepsilon = \pm 1$, multiplier à gauche une matrice C par $T_{i,j}(\varepsilon)T_{j,i}(-\varepsilon)T_{i,j}(\varepsilon)$ remplace L_i par εL_j, et L_j par $-\varepsilon L_i$. De même, multiplier à droite par $T_{i,j}(\varepsilon)T_{j,i}(-\varepsilon)T_{i,j}(\varepsilon)$ remplace C_i par $-\varepsilon C_j$ et C_j par εC_i.
En particulier, pour ramener le coefficient a_{ij} de C en position $(1,1)$, il suffit de multiplier à gauche par la matrice $T_{1,i}(1)T_{i,1}(-1)T_{1,i}(1)$ et à droite par la matrice $T_{1,j}(-1)T_{j,1}(1)T_{1,j}(-1)$.

(3) Les matrices de dilatation ne sont pas nécessaires si l'on ne souhaite pas normaliser les coefficients $a_1, \ldots, a_r$. Dans ce cas, ces derniers sont seulement uniques à association près. En effet, une légère modification des arguments de la démonstration précédente montrent que l'idéal engendré par les j-mineurs de C est l'idéal engendré par $a_1 \cdots a_j$, pour tout $j \in [\![0, r]\!]$.

(4) Ce théorème est encore vrai si A est seulement supposé principal, mais on a alors besoin de matrices supplémentaires, à savoir les matrices de Bézout. Ce type de matrices permet de faire apparaître un pgcd de deux éléments a et b dans une matrice, en s'appuyant sur l'existence d'une relation de Bézout. On démontre alors que l'algorithme s'arrête, en considérant la fonction

$$\begin{array}{rccc} \omega : & A \setminus \{0\} & \longrightarrow & \mathbb{N} \\ & a & \longmapsto & \displaystyle\sum_{\pi \in \mathscr{P}} v_\pi(a), \end{array}$$

et en remarquant que si $d \in A\setminus\{0\}$ est un diviseur strict de $a \in A\setminus\{0\}$, alors on a $\omega(d) < \omega(a)$. On perd néanmoins le côté algorithmique du théorème, puisque dans la pratique, on ne sait pas trouver une relation de Bézout dans un anneau principal quelconque. Le lecteur intéressé se référera à [9, Chapitre 8] pour plus de détails.

2.5. Définition. Soit A un anneau euclidien, soit $\mathscr{P}$ un s.c.r.i. de A fixé et soit $C \in \mathrm{M}_{m\times n}(A)$.

La matrice $E(a_1, \ldots, a_r)$ donnée par le théorème 2.3 s'appelle la *forme normale de Smith* de C (par rapport à $\mathscr{P}$).

Nous allons maintenant détailler deux exemples.

2.6. Exemple. Soit $A = \mathbb{Z}$ et $C = \begin{pmatrix} 10 & 14 \\ 6 & 7 \end{pmatrix} \in \mathrm{M}_2(\mathbb{Z})$.

On échange d'abord les lignes pour placer le terme de plus petite valeur absolue en position $(1,1)$. On obtient

$$C \sim \begin{pmatrix} 6 & 7 \\ 10 & 14 \end{pmatrix}.$$

L'opération $L_2 \longleftarrow L_2 - L_1$ donne

$$C \sim \begin{pmatrix} 6 & 1 \\ 10 & 4 \end{pmatrix}.$$

En échangeant les colonnes, on a

$$C \sim \begin{pmatrix} 1 & 6 \\ 4 & 10 \end{pmatrix}.$$

En utilisant des opérations élémentaires pour faire apparaître des 0, on obtient successivement

$$C \sim \begin{pmatrix} 1 & 0 \\ 4 & -14 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 \\ 0 & -14 \end{pmatrix}.$$

Enfin, en utilisant une matrice de dilatation, on a finalement

$$C \sim \begin{pmatrix} 1 & 0 \\ 0 & 14 \end{pmatrix}.$$

On obtient ainsi $r = 2, a_1 = 1$ et $a_2 = 14$. Cela pouvait d'ailleurs être obtenu sans opérations élémentaires en utilisant les formules explicites données par le théorème 2.3.

2.7. Exemple. Considérons la matrice

$$C = \begin{pmatrix} X^3 + X^2 - 4X + 2 & X^3 + 2X^2 - 3X \\ 4X^3 - X^2 - 6X + 3 & 4X^3 + 3X^2 - 7X \end{pmatrix} \in \mathrm{M}_2(\mathbb{R}[X]).$$

On choisit l'ensemble des polynômes unitaires, auquel on adjoint aussi le polynôme nul, comme s.c.r.i. de $\mathbb{R}[X]$.

On pourrait appliquer aveuglément l'algorithme, mais il est plus judicieux de remarquer que l'opération $C_2 \longleftarrow C_2 - C_1$ abaisse un peu le degré des coefficients de la seconde colonne. On obtient alors

$$C \sim \begin{pmatrix} X^3 + X^2 - 4X + 2 & X^2 + X - 2 \\ 4X^3 - X^2 - 6X + 3 & 4X^2 - X - 3 \end{pmatrix}.$$

Le polynôme $X^3 + X^2 - 4X + 2$ ne divise aucun élément de la première ligne ou de la première colonne. On peut donc utilisant des opérations élémentaires pour faire apparaître un pgcd de X^3+X^2-4X+2 et X^2+X-2, par exemple.

L'opération $L_1 \longleftarrow L_1 - XL_2$ donne

$$C \sim \begin{pmatrix} -2X + 2 & X^2 + X - 2 \\ -3X + 3 & 4X^2 - X - 3 \end{pmatrix}.$$

En divisant la première ligne par -2, on obtient

$$C \sim \begin{pmatrix} X-1 & -\frac{1}{2}(X^2+X-2) \\ -3X+3 & 4X^2-X-3 \end{pmatrix}.$$

Cette fois, $X-1$ divise déjà tous les coefficients de la première ligne et de la première colonne, et l'on peut passer l'étape suivante, c'est-à-dire faire apparaître des 0. En deux opérations élémentaires, on obtient finalement

$$C \sim \begin{pmatrix} X-1 & -\frac{1}{2}(X^2+X-2) \\ 0 & \frac{5}{2}(X^2-X) \end{pmatrix} \sim \begin{pmatrix} X-1 & 0 \\ 0 & \frac{5}{2}(X^2-X) \end{pmatrix}.$$

En multipliant la seconde colonne par $\frac{2}{5}$, on obtient

$$C \sim \begin{pmatrix} X-1 & 0 \\ 0 & X^2-X \end{pmatrix},$$

et l'on a fini, car $X-1$ divise X^2-X.

Nous verrons d'autres exemples plus loin, lorsque nous exposerons les diverses applications de ce théorème.

Pour conclure ce paragraphe, nous allons maintenant décrire des familles de générateurs de $\mathrm{GL}_n(A)$ et $\mathrm{SL}_n(A)$ lorsque A est un anneau euclidien.

2.8. Théorème. *Soit A un anneau euclidien. Alors, pour tout $n \geqslant 2$, le groupe $\mathrm{GL}_n(A)$ est engendré par chacune des familles suivantes :*

(1) *les matrices* $T_{i,j}^{(n)}(a)$, $S_{i,j}^{(n)}$, $D^{(n)}(u)$, $u \in A^\times$, $a \in A$, $i<j$, $i,j \in [\![1,n]\!]$;

(2) *les matrices* $T_{i,j}^{(n)}(a)$, $D^{(n)}(u)$, $u \in A^\times$, $i, j \in [\![1,n]\!]$.

Démonstration. Soit $C \in \mathrm{GL}_n(A)$. Puisque A est un anneau euclidien, l'algorithme de réduction permet d'obtenir $U, V \in \mathrm{GL}_n(A)$, toutes deux produit de matrices de permutation, de transvection ou de dilatation telles que

$$UCV \sim E(a_1, \ldots, a_r),$$

où $a_1, \ldots, a_r$ sont normalisés. Mais, les matrices U, C et V étant inversibles, la matrice $E(a_1, \ldots, a_r)$ l'est aussi. Cela implique que $r = n$ et que tous les a_i sont inversibles. Mais, les a_i étant normalisés, on a $a_i = 1$ pour tout entier $i \in [\![1,n]\!]$. Ainsi, on a en fait $UCV = \mathrm{I}_n$, et donc $C = U^{-1}V^{-1}$.

Pour démontrer (1), il reste donc à vérifier que l'on peut se limiter aux indices i, j tels que $i < j$ en ce qui concerne les matrices de transposition

et de transvection. Pour cela, il suffit de remarquer que si $i < j$, on a

$$S_{j,i}^{(n)} = S_{i,j}^{(n)} \quad \text{et} \quad S_{i,j}^{(n)} T_{i,j}^{(n)}(a) S_{i,j}^{(n)} = T_{j,i}^{(n)}(a).$$

Il reste à vérifier (2), ce qui provient de la remarque 2.4 (2). □

Avant de traiter le cas de $\mathrm{SL}_n(A)$, nous allons démontrer un lemme.

2.9. Lemme. *Soit A un anneau commutatif, soient $u_1, \ldots, u_n \in A^\times$ et soit* $D = \begin{pmatrix} u_1 & & \\ & \ddots & \\ & & u_n \end{pmatrix}$.

Alors, il existe U, $V \in \mathrm{SL}_n(A)$, toutes deux produit de matrices de transvection, telles que

$$UDV = \begin{pmatrix} \mathrm{I}_{n-1} & \\ & u_1 \cdots u_n \end{pmatrix}.$$

Démonstration. Nous affirmons que la sous-matrice $\begin{pmatrix} u_1 & 0 \\ 0 & u_2 \end{pmatrix}$ peut être ramenée à la matrice $\begin{pmatrix} 1 & 0 \\ 0 & u_1 u_2 \end{pmatrix}$, en manipulant seulement les deux premières lignes et les deux premières colonnes de D. Pour le voir, il suffit de faire les opérations suivantes :

$$L_1 \longleftarrow L_1 + L_2 \qquad \begin{pmatrix} u_1 & 0 \\ 0 & u_2 \end{pmatrix} \sim \begin{pmatrix} u_1 & u_2 \\ 0 & u_2 \end{pmatrix}$$

$$C_1 \longleftarrow C_1 + u_2^{-1}(1 - u_1) C_2 \qquad \begin{pmatrix} u_1 & u_2 \\ 0 & u_2 \end{pmatrix} \sim \begin{pmatrix} 1 & u_2 \\ 1 - u_1 & u_2 \end{pmatrix}$$

$$L_2 \longleftarrow L_2 + (u_1 - 1) L_1 \qquad \begin{pmatrix} 1 & u_2 \\ 1 - u_1 & u_2 \end{pmatrix} \sim \begin{pmatrix} 1 & u_2 \\ 0 & u_1 u_2 \end{pmatrix}$$

$$C_2 \longleftarrow C_2 - u_2 C_1 \qquad \begin{pmatrix} 1 & u_2 \\ 0 & u_1 u_2 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 \\ 0 & u_1 u_2 \end{pmatrix}.$$

Ainsi, il existe U', $xsV' \in \mathrm{SL}_n(A)$, toutes deux produits de matrices de transvection, telles que

$$U'DV' = \begin{pmatrix} 1 & & & & \\ & u_1 u_2 & & & \\ & & u_3 & & \\ & & & \ddots & \\ & & & & u_n \end{pmatrix}.$$

Il suffit alors de raisonner par récurrence pour obtenir le résultat souhaité. □

Passons maintenant au cas de $\mathrm{SL}_n(A)$. Pour tous $i, j \in [\![1, n]\!]$, avec $i \neq j$, posons

$$\hat{S}_{i,j}^{(n)} = E_{ij} - E_{ji} + \sum_{k \neq i,j} E_{kk}.$$

Autrement dit, $\hat{S}_{i,j}^{(n)}$ est la matrice

$$\begin{array}{c} \\ i \to \\ \\ \\ j \to \\ \\ \end{array} \begin{array}{c} \begin{array}{ccccccc} & & i & & j & & \\ & & \downarrow & & \downarrow & & \end{array} \\ \begin{pmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 1 & & & & \\ & & & 0 & & 1 & \\ & & & & 1 & & \\ & & & & & \ddots & \\ & & & & & & 1 \\ & & & -1 & & 0 & \\ & & & & & & 1 \\ & & & & & & & \ddots \\ & & & & & & & & 1 \end{pmatrix} \end{array} \in \mathrm{M}_n(A).$$

Il est clair que $\hat{S}_{i,j}^{(n)} \in \mathrm{SL}_n(A)$. Une telle matrice s'appelle une *matrice de transposition signée.*

Le lecteur vérifiera lui-même les points suivants :

(1) multiplier une matrice C à gauche par $\hat{S}_{i,j}^{(n)}$ revient à remplacer L_i par L_j, et L_j par $-L_i$;

(2) multiplier une matrice C à gauche par $\left(\hat{S}_{i,j}^{(n)}\right)^{-1}$ revient à remplacer L_i par $-L_j$, et L_j par L_i ;

(3) multiplier une matrice C à droite par $\hat{S}_{i,j}^{(n)}$ revient à remplacer C_i par $-C_j$, et C_j par C_i ;

(4) multiplier une matrice C à droite par $\left(\hat{S}_{i,j}^{(n)}\right)^{-1}$ revient à remplacer C_i par C_j, et C_j par $-C_i$.

Remarquons aussi que les matrices de transvection sont de déterminant 1. On obtient alors le résultat suivant.

2.10. Théorème. *Soit A un anneau euclidien. Alors, pour tout $n \geqslant 2$, le groupe $\mathrm{SL}_n(A)$ est engendré par chacune des familles suivantes :*

(1) *les matrices* $T_{i,j}^{(n)}(a)$, $\hat{S}_{i,j}^{(n)}$, $a \in A$, $i < j$, $i, j \in [\![1, n]\!]$;

(2) *les matrices* $T_{i,j}^{(n)}(a)$, $a \in A$, $i, j \in [\![1, n]\!]$.

Démonstration. Soit $C \in \mathrm{SL}_n(A)$. Tout d'abord, en adaptant la démonstration de l'algorithme (en remplaçant les matrices de transpositions par les matrices $\hat{S}_{i,j}^{(n)}$ et en supprimant l'usage des matrices de dilatation), on montre qu'il existe alors $U, V \in \mathrm{GL}_n(A)$, toutes deux produit de matrices

de transvection ou de matrices $\hat{S}_{i,j}^{(n)}$ telles que

$$UCV = E(a_1, \ldots, a_r),$$

où les $a_1, \cdots, a_r$ sont des éléments non nuls de A non nécessairement normalisés. Les matrices de transvection et les matrices $\hat{S}_{i,j}^{(n)}$ étant de déterminant 1, on a

$$U,\, V \in \mathrm{SL}_n(A).$$

Par suite, puisque $C \in \mathrm{SL}_n(A)$, on a aussi $E(a_1, \ldots, a_r) \in \mathrm{SL}_n(A)$. On a donc $r = n$ et $a_1 \cdots a_n = 1$. En particulier, tous les a_i sont inversibles. Le lemme précédent montre alors qu'il existe $U', V' \in \mathrm{SL}_n(A)$, toutes deux produit de matrices de transvection telles que

$$U' \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix} V' = \mathrm{I}_n.$$

Ainsi, $U'UCVV' = \mathrm{I}_n$, et donc $C = (U'U)^{-1}(VV')^{-1}$.

Pour démontrer (1), il reste donc à vérifier que l'on peut se limiter aux indices i, j tels que $i < j$ en ce qui concerne les matrices $\hat{S}_{i,j}^{(n)}$ et les matrices de transvection. Pour cela, il suffit de remarquer que si $i < j$, on a

$$\hat{S}_{i,j}^{(n)} = (\hat{S}_{j,i}^{(n)})^{-1},$$

$$\hat{S}_{i,j}^{(n)} T_{i,j}^{(n)}(a) \big(\hat{S}_{i,j}^{(n)}\big)^{-1} = T_{j,i}^{(n)}(a), \quad \text{si } j \neq n,$$

$$\hat{S}_{i,n}^{(n)} T_{i,n}^{(n)}(-a) \big(\hat{S}_{i,n}^{(n)}\big)^{-1} = T_{n,i}^{(n)}(a).$$

Enfin, le point (2) découle directement de la remarque 2.4 (2). □

En remarquant que l'on a

$$T_{i,j}^{(n)}(m) = \big(T_{i,j}^{(n)}(1)\big)^m, \quad \text{pour tout } m \in \mathbb{Z},$$

on obtient le corollaire suivant.

2.11. Corollaire. *Pour tout $n \geqslant 2$, le groupe $\mathrm{SL}_n(\mathbb{Z})$ est engendré par les matrices*

$$T_{i,j}^{(n)}(1) \quad \textit{et} \quad \hat{S}_{i,j}^{(n)}, \quad i < j,\ i, j \in [\![1, n]\!].$$

2.12. Remarque. En particulier, on obtient que $\mathrm{SL}_2(\mathbb{Z})$ est engendré par les matrices

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \text{et} \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix},$$

ce qui est bien connu.

Avant de pouvoir donner d'autres applications du théorème de réduction des matrices, il est commode d'introduire le langage des modules. Nous nous limiterons au strict minimum, et renvoyons le lecteur à [9] pour une introduction plus complète à ce sujet.

3. Modules sur un anneau

Commençons tout de suite par introduire la notion principale de ce paragraphe.

3.1. Définition. Soit A un anneau (non nécessairement commutatif).

Un *A-module* est un ensemble M muni d'une loi interne

$$\begin{aligned} M \times M &\longrightarrow M \\ (x, y) &\longmapsto x + y \end{aligned}$$

et d'une loi externe

$$\begin{aligned} A \times M &\longrightarrow M \\ (a, x) &\longmapsto a \cdot x \end{aligned}$$

appelée parfois *multiplication par un scalaire*, satisfaisant aux propriétés suivantes :

(1) l'ensemble M, muni de la loi « $+$ », est un groupe abélien ;

(2) on a $a \cdot (x + y) = a \cdot x + a \cdot y$ pour tous x, $y \in M$, et tout $a \in A$;

(3) on a $1_A \cdot x = x$ pour tout $x \in M$;

(4) on a $(a + b) \cdot x = a \cdot x + b \cdot x$ pour tout $x \in M$, et tous a, $b \in A$;

(5) on a $(ab) \cdot x = a \cdot (b \cdot x)$ pour tous a, $b \in A$, et tout $x \in M$.

Très grossièrement, un module est donc un « espace vectoriel sur un anneau ». La similitude de cette définition avec celle d'un espace vectoriel est néanmoins trompeuse. Nous verrons que si A est un anneau arbitraire, les propriétés des espaces vectoriels ne se transposent pas au cas des modules. Donnons maintenant quelques exemples.

3.2. Exemples

(1) Si A est un corps K, un module n'est rien d'autre qu'un K-espace vectoriel.

(2) La loi de multiplication interne fait de tout anneau A un A-module.

(3) Un groupe abélien G est un $\mathbb{Z}$-module pour la loi externe

$$\begin{aligned} \mathbb{Z} \times G &\longrightarrow G \\ (m, g) &\longmapsto m \cdot g, \end{aligned}$$

$$\text{où } m \cdot g = \begin{cases} \underbrace{g + \cdots + g}_{m \text{ fois}} & \text{si} \quad m > 0 \\ 0 & \text{si} \quad m = 0 \\ \underbrace{(-g) + \cdots + (-g)}_{-m \text{ fois}} & \text{si} \quad m < 0. \end{cases}$$

(4) Pour tout $n \geqslant 1$, le groupe abélien A^n est un A-module pour la loi externe

$$\begin{aligned} A \times A^n &\longrightarrow A^n \\ \left(a, \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) &\longmapsto \begin{pmatrix} ax_1 \\ \vdots \\ ax_n \end{pmatrix}. \end{aligned}$$

(5) Plus généralement, si $M_1, \ldots, M_n$ sont des A-modules, le groupe abélien $M_1 \times \cdots \times M_n$ est un A-module pour la loi externe

$$\begin{aligned} A \times (M_1 \times \cdots \times M_n) &\longrightarrow M_1 \times \cdots \times M_n \\ \big(a, (x_1, \ldots, x_n)\big) &\longmapsto (a \cdot x_1, \ldots, a \cdot x_n). \end{aligned}$$

Les démonstrations sont laissées au lecteur à titre d'exercice.

Nous continuons en donnant quelques conséquences de la structure de module, qui se démontrent exactement de la même manière que dans le cas des espaces vectoriels.

3.3. Lemme. *Soit M un A-module. Alors, on a les relations suivantes :*

(1) *pour tout $x \in M$, on a $0_A \cdot x = 0_M$;*

(2) *pour tout $a \in A$, on a $a \cdot 0_M = 0_M$;*

(3) *pour tout $x \in M$, et tout $a \in A$, on a $(-a) \cdot x = a \cdot (-x) = -(a \cdot x)$. En particulier, $(-1_A) \cdot x = -x$, pour tout $x \in M$.*

De même que l'on peut définir la notion de sous-groupe ou de sous-espace vectoriel, nous avons une notion de sous-module.

3.4. Définition. Soit M un A-module. Un *sous-module* de M est un sous-ensemble non vide N de M vérifiant les propriétés suivantes :

(1) pour tous $x, y \in N$, on a $x + y \in N$;

(2) pour tout $x \in N$, et tout $a \in A$, on a $a \cdot x \in N$.

De manière équivalente, un sous-module de M est un sous-ensemble non vide N de M vérifiant

$$a{\cdot}x + b{\cdot}y \in N, \quad \text{pour tous } x,\, y \in N \text{ et tous } a,\, b \in A.$$

Donnons maintenant des exemples de sous-modules.

3.5. Exemples

(1) Si A est un corps K, on retrouve la notion de sous-espace vectoriel.

(2) Considérons A comme A-module. Alors, un sous-module est simplement un idéal à gauche de A.

(3) Si M est un A-module, l'ensemble $\{0_M\}$ est un sous-module de M, et souvent noté simplement 0 par abus de notation. De même, M est un sous-module de M.

(4) Si G est un groupe abélien, vu comme $\mathbb{Z}$-module, alors un sous-module n'est rien d'autre qu'un sous-groupe de G.

Comme dans le cas des espaces vectoriels, un sous-module N d'un module M est un A-module pour les lois induites. De même, l'intersection d'une famille de sous-modules de M est un sous-module de M. On peut donc poser la définition suivante.

3.6. Définition. Soit M un A-module, et soit $\mathscr{P}$ une partie de M. Le *sous-module de M engendré par $\mathscr{P}$*, noté $\langle \mathscr{P} \rangle$, est l'intersection de tous les sous-modules de M contenant $\mathscr{P}$.

Comme en algèbre linéaire, on a une description précise du sous-module engendré par $\mathscr{P}$.

3.7. Lemme. *Soit M un A-module. Pour tout partie $\mathscr{P}$ de M, on a*

$$\langle \mathscr{P} \rangle = \Big\{ \sum_{i=1}^{n} a_i{\cdot}x_i \mid a_i \in A,\ x_i \in \mathscr{P},\ n \geqslant 0 \Big\}.$$

3.8. Remarque. Si $\mathscr{P} = \{x_1, \cdots, x_m\}$ est fini, on a

$$\langle \mathscr{P} \rangle = \Big\{ \sum_{i=1}^{m} a_i{\cdot}x_i \mid a_i \in A \Big\},$$

et l'on note parfois $\langle \mathscr{P} \rangle$ par $A{\cdot}x_1 + \cdots + A{\cdot}x_m$.

3.9. Définition. Un élément de $\langle \mathscr{P} \rangle$ est appelé une *combinaison A-linéaire* d'éléments de $\mathscr{P}$. On dira « combinaison linéaire » lorsque A est clairement défini par le contexte.

3.10. Définition. Soit A un anneau, et soit M un A-module. On dit qu'une partie $\mathscr{P}$ de M est *A-génératrice* si $M = \langle \mathscr{P} \rangle$, c'est-à-dire si tout élément de M s'écrit comme combinaison linéaire d'éléments de $\mathscr{P}$. On dit aussi que M est engendré par $\mathscr{P}$, ou que $\mathscr{P}$ engendre M. On abrégera souvent en disant « famille génératrice » lorsque A est clairement défini par le contexte.

On dira que M est *de type fini* s'il est engendré par un nombre fini d'éléments.

Comme en algèbre linéaire, on a la notion d'application A-linéaire.

3.11. Définition. Une *application A-linéaire* est une application

$$f : M \longrightarrow N$$

satisfaisant les propriétés suivantes :

(1) pour tous $x, y \in M$, on a $f(x+y) = f(x) + f(y)$;

(2) pour tout $a \in A$, et tout $x \in X$, on a $f(a \cdot x) = a \cdot f(x)$.

Lorsque le contexte est clair, on parlera simplement d'« application linéaire ». On vérifie sans peine que la somme, la différence et la composée d'applications linéaires est encore une application linéaire. En particulier, l'ensemble des applications linéaires de M vers N est un groupe additif, noté $\mathrm{Hom}_A(M, N)$. Si $M = N$, on parle aussi d'*endomorphisme de M*, et l'on note $\mathrm{End}_A(M)$ le groupe $\mathrm{Hom}_A(M, M)$. C'est un anneau pour l'addition et la composition des endomorphismes.

Un *isomorphisme* de A-modules est une application A-linéaire bijective. L'application inverse f^{-1} est alors aussi linéaire. L'ensemble des isomorphismes de M sur N est noté $\mathrm{Iso}_A(M, N)$. Un *automorphisme* de M est un isomorphisme de M sur lui-même. On vérifie que l'ensemble des automorphismes de M est un groupe pour la composition, appelé *groupe linéaire* du A-module M, et que l'on note $\mathrm{GL}_A(M)$.

3.12. Remarque. Si A est un corps, on retrouve simplement la notion d'application linéaire, d'endomorphisme, d'isomorphisme et d'automorphisme d'espaces vectoriels.

3.13. Remarque. Si A est **commutatif**, le groupe abélien $\mathrm{Hom}_A(M, N)$ possède une structure naturelle de A-module.

En effet, si $f \in \mathrm{Hom}_A(M,N)$ et $a \in A$, on vérifie que l'application

$$\begin{aligned} a{\cdot}f : M &\longrightarrow N \\ x &\longmapsto a{\cdot}f(x) \end{aligned}$$

est A-linéaire (cela utilise la commutativité de A). La loi externe

$$\begin{aligned} A \times \mathrm{Hom}_A(M,N) &\longrightarrow \mathrm{Hom}_A(M,N) \\ (a,f) &\longmapsto a{\cdot}f \end{aligned}$$

induit alors sur $\mathrm{Hom}_A(M,N)$ une structure de A-module. Les détails sont laissés au lecteur.

3.14. Remarque. Si $f : M \longrightarrow N$, on a nécessairement $f(0) = 0$.

En effet, on a $f(0) = f(0+0) = f(0)+f(0)$, d'où la conclusion en simplifiant par $f(0)$.

Comme en algèbre linéaire, on a le résultat suivant.

3.15. Lemme. *Soit $f : M \longrightarrow N$ une application linéaire. Alors :*

(1) *pour tout sous-module M' de M, $f(M')$ est un sous-module de N ;*

(2) *pour tout sous-module N' de N, $f^{-1}(N')$ est un sous-module de M.*

En particulier, les ensembles $\mathrm{Ker}(f) = f^{-1}(\{0\})$ *et* $\mathrm{Im}(f) = f(M)$ *sont des sous-modules de M et N respectivement.*

3.16. Définition. Les A-modules $\mathrm{Ker}(f)$ et $\mathrm{Im}(f)$ sont respectivement appelés *noyau* et *image* de f de l'application linéaire $f : M \longrightarrow N$.

On vérifie qu'une application linéaire $f : M \longrightarrow N$ est injective si, et seulement si, $\mathrm{Ker}(f) = 0$.

On s'intéresse maintenant à la notion de module libre, qui généralise plutôt bien la notion d'espace vectoriel. Comme on le verra un peu plus loin, tous les A-modules ne sont pas libres. On commence par définir les notions de famille libre et de base.

3.17. Définition. Soit A un anneau, et soit M un A-module.

Une famille $(x_i)_{i\in I}$ d'éléments de M est dite *A-libre* si pour tout sous-ensemble fini $J \subset I$, et tout $a_j \in A$, $j \in J$, on a

$$\sum_{j\in J} a_j{\cdot}x_j = 0 \quad \Longrightarrow \quad a_j = 0, \text{ pour tout } j \in J.$$

Une famille qui n'est pas A-libre est dite *A-liée*. On dira aussi famille « libre » ou « liée », lorsque A est clairement défini par le contexte.

Une *A-base* est une famille $(e_i)_{i\in I}$ d'éléments de M qui à la fois libre et génératrice. Comme toujours, on parlera simplement de « base » lorsque le contexte est clair. De manière équivalente, $(e_i)_{i\in I}$ est une base de M si tout élément $x \in M$ s'écrit de manière unique sous la forme

$$x = \sum_{i\in I} a_i \cdot e_i, \quad a_i \in A \text{ presque tous nuls.}$$

On dit que M est un A-module *libre* s'il possède une base.

Remarquons que, au vu des définitions, l'ensemble vide est une base du module nul.

3.18. Exemples

(1) Pour tout $n \geqslant 1$, le A-module A^n est libre. En effet, pour tout $i \in I$, soit ε_i l'élément de A^n dont toutes les coordonnées sont nulles, sauf celle qui correspond à l'indice i, qui vaut 1. Il est alors facile de voir que la famille $(\varepsilon_1, \ldots, \varepsilon_n)$ est une base de A^n, appelée *base canonique.*

(2) Un K-espace vectoriel de dimension finie est un K-module libre, et tout sous-module d'un K-module libre est libre.

(3) Soit maintenant A un anneau intègre, vu comme A-module, et soit I un sous-module de A, c'est-à-dire un idéal de A. Alors, I est libre si et seulement si, il est principal.

Supposons en effet que I soit principal. Si I est nul, alors il est libre de base $\emptyset$. Sinon $I = (a)$, $a \in A \setminus \{0\}$, et a est une base de I. En effet, a engendre I, et de plus a est libre puisque par intégrité de A, on a

$$a' \cdot a = a' \cdot a = 0 \quad \Longrightarrow \quad a' = 0.$$

Réciproquement, supposons que I soit un A-module libre. Remarquons que deux éléments a, $b \in I$ distincts sont A-liés (et donc en particulier un des deux éléments est non nul), puisque l'on a

$$a \cdot b - b \cdot a = ab - ba = 0.$$

Ainsi, une base de I possède au plus un élément. Comme une base est génératrice, on a soit $I = (0)$ (si la base est vide), soit $I = (a)$, pour un élément $a \in A \setminus \{0\}$. Ainsi, I est principal dans tous les cas.

En particulier, si I est un idéal non principal, il n'est pas libre, bien qu'il soit sous-module du A-module libre A. Pour avoir un exemple concret, il suffit de prendre $A = \mathbb{Z}[X]$ et $I = (2, X)$.

Attention ! La plupart des résultats valables pour les espaces vectoriels ne le sont plus pour les A-modules libres (cf. exercice 2).

Nous allons maintenant nous concentrer sur les modules libres de type fini.

3.19. Lemme. *Soit A un anneau, et soit M un A-module libre de type fini. Alors, toutes les bases de M ont un cardinal fini.*

Démonstration. Soit $(x_1, \ldots, x_n)$ une famille génératrice et soit $(e_i)_{i\in I}$ une base de M. Par hypothèse, chaque x_i est une combinaison linéaire finie des e_i. Il existe donc un sous-ensemble fini d'indices $J \subset I$ tel que pour tout entier $m \in [\![1, n]\!]$, x_m est une combinaison linéaire des $(e_j)_{j\in J}$. Tout élément de M étant une combinaison linéaire de $x_1, \ldots, x_n$, il s'ensuit que la famille $(e_j)_{j\in J}$ est également génératrice. Supposons maintenant que $J \neq I$, et soit $i_0 \in I \setminus J$. Puisque $(e_j)_{j\in J}$ est génératrice, e_{i_0} est combinaison linéaire des e_j, $j \in J$. En particulier, la famille $(e_i)_{i\in I}$ est liée, d'où une contradiction. Ainsi, $I = J$, et donc toute base de M est de cardinal fini.□

Comme dans le cas des espaces vectoriels, les applications linéaires d'un module libre de type fini vers un module quelconque sont entièrement déterminées par le choix des images des éléments d'une base.

3.20. Lemme. *Soit M un A-module libre de type fini, soit $(e_1, \ldots, e_n)$ une base de M et soit M' un A-module. Pour toute famille $(y_1, \ldots, y_n)$ d'éléments de M', il existe une unique application linéaire $f : M \longrightarrow M'$ telle que*

$$f(e_i) = y_i, \qquad \textit{pour tout } i \in [\![1, n]\!].$$

Elle est donnée par

$$f\Big(\sum_{i=1}^{n} a_i \cdot e_i\Big) = \sum_{i=1}^{n} a_i \cdot y_i, \qquad \textit{pour tout } a_i \in A.$$

3.21. Corollaire. *Soit M un A-module. Alors, M est un A-module libre de type fini si, et seulement si, $M \simeq A^n$ pour un certain entier*[3] *$n \geqslant 0$.*

Démonstration. Si M est un A-module libre de type fini, alors il admet une base finie $\mathbf{e} = (e_1, \ldots, e_n)$ par le lemme 3.19. On vérifie alors que l'unique application linéaire $f : A^n \longrightarrow M$ tell que $f(\varepsilon_i) = e_i$ pour tout $i \in [\![1, n]\!]$ est un isomorphisme. Inversement, si $f : A^n \xrightarrow{\sim} M$ est un isomorphisme, on vérifie que l'image de la base canonique de A^n est une base de M. □

Soient M et M' deux A-modules libres de type fini, de bases $\mathbf{e} = (e_1, \ldots, e_n)$ et $\mathbf{e}' = (e'_1, \ldots, e'_m)$ (le fait que ces bases ont un nombre fini d'éléments provient du lemme 3.19). Le résultat précédent implique alors qu'une application linéaire $f : M \longrightarrow M'$ est entièrement déterminée par les coordonnées

3. On adopte la convention $A^0 = (0)$.

de $f(e_1), \ldots, f(e_n)$ dans la base $\mathbf{e}'$, tout comme dans le cas des espaces vectoriels. On définit alors la matrice représentative $\mathrm{Mat}(f\,; \mathbf{e}, \mathbf{e}')$ de la manière habituelle.

Lorsque A est **commutatif**, cette matrice a les propriétés usuelles. On a alors le théorème suivant :

3.22. Théorème. *Soit A un anneau commutatif. Alors, toutes les bases d'un A-module libre de type fini ont même cardinal.*

Démonstration. Soit M un A-module libre de type fini.

Soient $\mathbf{e} = (e_1, \ldots, e_n)$, $\mathbf{f} = (f_1, \ldots, f_m)$ deux bases de M. Enfin, soient

$$P = \mathrm{Mat}(\mathrm{Id}_M; \mathbf{e}, \mathbf{f}) \in \mathrm{M}_{m\times n}(A), \quad Q = \mathrm{Mat}(\mathrm{Id}_M; \mathbf{f}, \mathbf{e}) \in \mathrm{M}_{n\times m}(A).$$

Alors, on a

$$PQ = \mathrm{I}_m \quad \text{et} \quad QP = \mathrm{I}_n.$$

Soit $\mathfrak{m}$ un idéal maximal de A, et soit $K = A/\mathfrak{m}$. C'est donc un corps.

Pour toute matrice $C = (a_{ij}) \in \mathrm{M}_{p\times q}(A)$ à coefficients dans A, notons alors $\overline{C} = (\overline{a}_{ij}) \in \mathrm{M}_{p\times q}(K)$. On vérifie sans mal que l'on a

$$\overline{P}\,\overline{Q} = \overline{PQ} = \overline{\mathrm{I}}_m = \mathrm{I}_m \in \mathrm{M}_m(K),$$

et de même $\overline{Q}\,\overline{P} = \mathrm{I}_n \in \mathrm{M}_n(K)$. Ainsi, $\overline{P}$ est une matrice inversible à coefficients dans un corps. Elle est donc carrée. Par conséquent, on a $n = m$, ce qui achève la démonstration. □

3.23. Remarques

(1) Ce résultat est faux si A n'est pas commutatif. Il existe même des anneaux A tel que le A-module A ait une base de cardinal n pour tout entier $n \geqslant 1$ (cf. exercice 3).

(2) En appliquant ce résultat à $A = \mathbb{Z}$, on retrouve le théorème XII-4.5.

3.24. Définition. Soit A un anneau commutatif, et soit M un A-module libre de type fini. Le cardinal d'une base de M est appelé le *rang* de M, et est noté $\mathrm{rg}_A(M)$.

Par définition, $\mathrm{rg}_A(A^n) = n$, pour tout $n \geqslant 0$.

On a alors le résultat suivant.

3.25. Théorème. *Soit A un anneau commutatif. Alors, les A-modules libres de type fini sont classés à isomorphisme près par leur rang.*

4. Théorème de la base adaptée et applications

On suppose dorénavant que A est un anneau euclidien.

4.1. Lemme. *Soit A un anneau euclidien. Tout sous-module d'un module libre de rang n est de type fini, engendré par au plus n éléments.*

Démonstration. On procède par récurrence sur n. Si $n = 0$, il n'y a rien à faire. Supposons que le résultat soit démontré au rang n, et démontrons-le au rang $n+1$.

Fixons une base $\mathbf{e} = (e_1, \ldots, e_{n+1})$ de M. Soit $(x_j)_{i \in J}$ une famille génératrice de N. On a donc

$$x_j = \sum_{i=1}^{n+1} a_{ij} \cdot e_i, \quad \text{pour tout } j \in J.$$

Soit $\mathfrak{a}$ l'idéal engendré par les $a_{n+1\ j}$, $j \in J$. Comme A est euclidien, donc principal, on a $\mathfrak{a} = (a)$. On peut écrire

$$a = \sum_{j \in J} \lambda_j a_{n+1\ j}, \quad \lambda_j \in A,$$

où les λ_j sont presque tous nuls. On pose

$$y_0 = \sum_{j \in J} \lambda_j \cdot x_j \in N.$$

Alors, la $(n+1)$-ième coordonnée de y_0 est a. On peut aussi écrire

$$a_{n+1\ j} = \mu_j a,$$

et ainsi le sous-module N' engendré par les $x_j - \mu_j y_0$, $j \in J$ est un sous-module du module M' engendré par $e_1, \ldots, e_n$. Alors, M' est un module libre de rang n. Par hypothèse, N' est engendré par

$$y_1, \ldots, y_m \in N' \subset N,$$

avec $m \leqslant n$. Mais alors, on en déduit que x_j est combinaison linéaire des éléments $y_0, y_1, \ldots, y_m$. On en déduit aisément que N est engendré par les éléments $y_0, \ldots, y_m$. Cela achève la récurrence (puisque $m + 1 \leqslant n + 1$). □

4.2. Remarques

(1) Un examen de la démonstration précédente montre en fait que ce résultat est encore vrai si A est seulement supposé principal. En revanche, il est devient faux pour un anneau quelconque. Par exemple, si $A = \mathbb{C}[X_1, \ldots, X_n, \ldots]$, alors A est libre de rang 1, mais l'idéal

$$I = (X_1, \ldots, X_n, \ldots)$$

est un sous-module de A qui ne peut pas être engendré par un nombre fini d'éléments (cf. exemple XXI-3.15 (2)).

(2) Si $A = \mathbb{Z}$, on retrouve le résultat du corollaire XII-4.11.

Dans tout ce qui suit, on suppose fixé une fois pour toutes un s.c.r.i. de A. On a le théorème suivant, qui précise énormément le lemme précédent.

4.3. Théorème (de la base adaptée). *Soit A un anneau euclidien. Soit M un A-module libre de rang m, et soit N un sous-module de M. Alors, il existe une base $(e_1, \dots, e_m)$ de M, un entier $r \geqslant 0$ et des éléments normalisés $a_1, \dots, a_r \in A$ non nuls vérifiant $a_1 \mid \cdots \mid a_r$, et tels que la famille $(a_1 \cdot e_1, \dots, a_r \cdot e_r)$ soit une base de N. En particulier, N est un module libre de rang $r \leqslant m$. De plus, l'entier r et les éléments $a_1, \dots, a_r$ sont uniques.*

Démonstration. Par le lemme précédent, N est de type fini, engendré par un nombre fini d'éléments $x_1, \dots, x_n$. Soit $(\varepsilon_1, \dots, \varepsilon_m)$ une base de M. Supposons que $N \neq 0$ (sinon il n'y a rien à faire). Écrivons

$$x_j = \sum_{i=1}^{m} a_{ij} \cdot \varepsilon_i, \quad \text{pour tout } j \in [\![1, n]\!]$$

et posons $C = (a_{ij}) \in \mathrm{M}_{m \times n}(A)$. D'après le théorème 2.3, il existe deux matrices $U \in \mathrm{GL}_m(A)$ et $V \in \mathrm{GL}_n(A)$ telles que

$$UCV = E(a_1, \dots, a_r),$$

avec $0 \leqslant r \leqslant \min(m, n) \leqslant m$ et $a_1, \dots, a_r$ vérifiant les conditions du théorème. Soient $e_1, \dots, e_m$ les vecteurs de M dont les vecteurs des coordonnées dans la base $(\varepsilon_1, \dots, \varepsilon_m)$ sont donnés par les colonnes de U^{-1}. Alors, $(e_1, \dots, e_m)$ est une base de M, car U^{-1} est inversible. Vérifions maintenant que $(a_1 \cdot e_1, \dots, a_r \cdot e_r)$ est une base de N. Remarquons que les r premières colonnes de la matrice $U^{-1}E(a_1, \dots, a_r)$ sont exactement les éléments $a_1 \cdot e_1, \dots, a_r \cdot e_r$ écrits dans la base $\varepsilon_1, \dots, \varepsilon_m$, et que les $n - r$ autres colonnes sont nulles. Par ailleurs, il est facile de voir que les colonnes de la matrice CV sont des combinaisons linéaires des colonnes de C (cela provient de la définition de la multiplication matricielle).

Puisque $U^{-1}E(a_1, \dots, a_r) = CV$, on obtient que $a_1 \cdot e_1, \dots, a_r \cdot e_r$ sont des combinaisons linéaires des $x_1, \dots, x_n$. On a donc

$$a_1 \cdot e_1, \ \dots, \ a_r \cdot e_r \in N.$$

Par ailleurs, les colonnes de la matrice $C = U^{-1}E(a_1, \dots, a_r)V^{-1}$ sont des combinaisons linéaires des colonnes de $U^{-1}E(a_1, \dots, a_r)$. Ainsi, les éléments $x_1, \dots, x_n$ sont combinaisons linéaires de $a_1 \cdot e_1, \dots, a_r \cdot e_r$. Il s'ensuit

que la famille $(a_1 \cdot e_1, \ldots, a_r \cdot e_r)$ engendre N. Puisque $(e_1, \ldots, e_m)$ est une base de M, les éléments $e_1, \ldots, e_r$ sont linéairement indépendants, et par conséquent la famille $(a_1 \cdot e_1, \ldots, a_r \cdot e_r)$ est aussi libre (car A est intègre). On a donc démontré que $(a_1 \cdot e_1, \ldots, a_r \cdot e_r)$ est une base de N.

Soient $\mathbf{e} = (e_1, \ldots, e_m)$ et $\mathbf{e}' = (e'_1, \ldots, e'_m)$ deux bases de M, et des éléments $a_1, \ldots, a_r, a'_1, \ldots, a'_s \in A$ vérifiant les conditions du théorème. Puisque toutes les bases d'un A-module libre ont même cardinal d'après le théorème 3.22, on a $r = s$. Puisque $\mathbf{e}$ et $\mathbf{e}'$ sont deux bases de M, on peut écrire

$$e_j = \sum_{i=1}^{m} u_{ij} \cdot e'_i, \quad \text{pour tout } j \in [\![1, m]\!].$$

Remarquons que la matrice $U = (u_{ij}) \in \mathrm{M}_m(A)$ est inversible par définition. Les égalités précédentes entraînent que

$$a_j \cdot e_j = \sum_{i=1}^{m} (a_j u_{ij}) \cdot e'_i, \quad \text{pour tout } j \in [\![1, r]\!].$$

Par ailleurs, $(a_1 \cdot e_1, \ldots, a_r \cdot e_r)$ et $(a'_1 \cdot e'_1, \ldots, a'_r \cdot e'_r)$ étant deux bases de N, on peut écrire

$$a_j \cdot e_j = \sum_{i=1}^{r} v_{ij} \cdot (a'_i \cdot e'_i) = \sum_{i=1}^{r} (v_{ij} a'_i) \cdot e'_i, \quad \text{pour tout } j \in [\![1, r]\!].$$

Remarquons que la matrice $V = (v_{ij}) \in \mathrm{M}_r(A)$ est aussi inversible. Par unicité de l'écriture dans une base, on a alors

$$a_j u_{ij} = v_{ij} a'_i, \quad \text{pour tous } i, j \in [\![1, r]\!]$$

et également

$$a_j u_{ij} = 0, \quad \text{pour tout } i \in [\![r+1, m]\!] \text{ et tout } j \in [\![1, r]\!].$$

Puisque les éléments $a_1, \ldots, a_r$ sont non nuls et que A est intègre, on a donc

$$u_{ij} = 0, \quad \text{pour tout } i \in [\![r+1, m]\!] \text{ et tout } j \in [\![1, r]\!].$$

Ainsi, U est de la forme

$$U = \begin{pmatrix} U_1 & U_2 \\ \mathrm{O}_{m-r,m-r} & U_3 \end{pmatrix}.$$

En utilisant le calcul matriciel par blocs, ou le fait que l'on a

$$\det(U) = \det(U_1) \det(U_3),$$

on voit que l'inversibilité de U implique celle de U_1. Les égalités

$$a_j u_{ij} = v_{ij} a'_i, \quad \text{pour tous } i, j \in [\![1, r]\!],$$

se traduisent facilement matriciellement par

$$U_1 E(a_1, \ldots, a_r) = E(a'_1, \cdots, a'_r)V,$$

où ici les matrices $E(a_1, \ldots, a_r)$ et $E(a'_1, \cdots, a'_r)$ sont des matrices carrées de taille r. Comme U_1 et V sont inversibles, on en déduit

$$E(a_1, \ldots, a_r) \sim E(a'_1, \cdots, a'_r),$$

et l'on conclut en utilisant la partie unicité du théorème 2.3. □

4.4. Remarque. Soit M un A-module libre, et soit N un sous-module. Pour calculer une base de M adaptée à N, on choisit une base de M, puis on écrit la matrice C des coordonnées d'une famille de générateurs de N dans cette base.

On applique l'algorithme de réduction pour trouver deux matrices inversibles U et V telles que

$$UCV = E(a_1, \ldots, a_r).$$

En fait, il suffit de calculer U^{-1} en même temps que l'on applique l'algorithme. Les colonnes de U^{-1} donnent les coordonnées de $e_1, \ldots, e_m$ dans la base de M fixée, et $(a_1 \cdot e_1, \ldots, a_r \cdot e_r)$ est une base de N.

Intéressons-nous maintenant au problème suivant : étant donnés une matrice C de $\mathrm{M}_{m\times n}(A)$ et $Y \in A^m$, quelles sont les conditions nécessaires et suffisantes pour que le système $CX = Y$ ait une solution $X \in A^n$? Dans ce cas, quel en est l'ensemble des solutions ?

4.5. Définition. Si $C \in \mathrm{M}_{m\times n}(A)$, on pose

$$\mathrm{Ker}(C) = \{X \in A^n \mid CX = 0\} \quad \text{et} \quad \mathrm{Im}(C) = \{CX \mid C \in A^n\}.$$

Alors, $\mathrm{Ker}(C)$ et $\mathrm{Im}(C)$ sont des sous-modules de A^n et A^m respectivement, appelé *noyau* et *image* de C. Ce sont en fait le noyau et l'image de l'application linéaire

$$\begin{aligned} f_C : A^n &\longrightarrow A^m \\ X &\longmapsto CX. \end{aligned}$$

Remarquons que $\mathrm{Im}(C)$ est aussi le sous-module de A^n engendré par les colonnes de C.

Clairement, le système $CX = Y$ admet au moins une solution si, et seulement si, $Y \in \mathrm{Im}(C)$. Si X_0 est une solution particulière du système, l'ensemble des solutions est alors $X_0 + \mathrm{Ker}(C)$.

Nous allons à présent étudier de plus près les A-modules $\mathrm{Ker}(C)$ et $\mathrm{Im}(C)$.

4.6. Proposition. *Soit A un anneau euclidien, et soit $C \in \mathrm{M}_{m\times n}(A)$ une matrice de rang r (comme matrice à coefficients dans K_A). Alors,* $\mathrm{Im}(C)$ *est un A-module libre de rang r. De plus, si $E(a_1,\dots,a_r)$ est la forme normale de Smith de C, et si*

$$UCV = E(a_1,\dots,a_r), \qquad \text{avec } U \in \mathrm{GL}_m(A),\ V \in \mathrm{GL}_n(A),$$

alors une base de $\mathrm{Im}(C)$ *est donnée par $a_1{\cdot}Y_1,\dots,a_r{\cdot}Y_r$, où $Y_1,\dots,Y_r$ sont les r premières colonnes de U^{-1}.*

Enfin, un vecteur $Y \in A^m$ est dans $\mathrm{Im}(C)$ *si, et seulement si, il existe des éléments $z_1,\dots,z_r \in A$ tels que*

$$UY = \begin{pmatrix} a_1z_1 \\ \vdots \\ a_rz_r \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

Démonstration. Gardons les notations de l'énoncé.

On a $C = U^{-1}E(a_1,\dots,a_r)V^{-1}$, et comme V est inversible, on en déduit

$$\mathrm{Im}(C) = \mathrm{Im}\big(U^{-1}E(a_1,\dots,a_r)\big).$$

Or, les r premières colonnes de $U^{-1}E(a_1,\dots,a_r)$ sont $a_1{\cdot}Y_1,\dots,a_r{\cdot}Y_r$ et les $n-r$ autres sont nulles. Ainsi, $\mathrm{Im}(C)$ est engendré par $a_1Y_1,\dots,a_rY_r$. Comme la matrice U est inversible, la famille $(Y_1,\dots,Y_r)$ est K_A-libre, donc aussi A-libre. Comme $a_1,\dots,a_r$ sont non nuls et l'anneau A est intègre, on en déduit que la famille $(a_1{\cdot}Y_1,\dots,a_r{\cdot}Y_r)$ est aussi A-libre. C'est donc une base de $\mathrm{Im}(C)$.

Remarquons maintenant que $Y \in \mathrm{Im}(C) = \mathrm{Im}\big(U^{-1}E(a_1,\dots,a_r)\big)$ si, et seulement si, $UY \in \mathrm{Im}\big(E(a_1,\dots,a_r)\big)$. Or, l'image de $E(a_1,\dots,a_r)$ est l'ensemble des vecteurs de A^m de la forme

$$\begin{pmatrix} a_1z_1 \\ \vdots \\ a_rz_r \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \text{avec } z_1,\dots,z_r \in A.$$

Cela achève la démonstration. □

4.7. Remarque. Si $UY = \begin{pmatrix} a_1 z_1 \\ \vdots \\ a_r z_r \\ 0 \\ \vdots \\ 0 \end{pmatrix}$, alors une solution particulière X_0 du système $CX = Y$ est donnée par [4]

$$X_0 = V \begin{pmatrix} z_1 \\ \vdots \\ z_r \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

En effet, on a alors

$$CX_0 = U^{-1}E(a_1, \ldots, a_r) \begin{pmatrix} z_1 \\ \vdots \\ z_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} = U^{-1} \begin{pmatrix} a_1 z_1 \\ \vdots \\ a_r z_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} = U^{-1}UY = Y.$$

Intéressons-nous maintenant à $\mathrm{Ker}(C)$.

4.8. Proposition. *Soit A un anneau euclidien, et soit $C \in \mathrm{M}_{m\times n}(A)$ une matrice de rang r (comme matrice à coefficients dans K_A). Alors,* $\mathrm{Ker}(C)$ *est un A-module libre de rang $n - r$.*

De plus, si l'on a

$$UCV = E(a_1, \ldots, a_r), \qquad \text{avec } U \in \mathrm{GL}_m(A), \quad V \in \mathrm{GL}_n(A),$$

alors une base de $\mathrm{Ker}(C)$ *est donnée par les $n - r$ dernières colonnes de V.*

Démonstration. Si $X \in A^n$, on a

$$\begin{aligned} CX = 0 &\iff U^{-1}E(a_1, \ldots, a_r)V^{-1}X = 0 \\ &\iff E(a_1, \ldots, a_r)V^{-1}X = 0. \end{aligned}$$

Ainsi, $X \in \mathrm{Ker}(C)$ si, et seulement si, $V^{-1}X \in \mathrm{Ker}\big(E(a_1, \ldots, a_r)\big)$. Or, il est facile de voir que le noyau de $E(a_1, \ldots, a_r)$ est le A-module libre dont une base est donnée par les $n - r$ derniers vecteurs de la base canonique de A^n. Le résultat en découle. □

Les résultats précédents nous donnent une description complète des solutions d'une équation linéaire à coefficients dans A.

4. Attention ! Malgré les apparences, les vecteurs UY et X_0 n'ont pas la même taille.

4.9. Corollaire. *Soit A un anneau euclidien. Soient $a_1, \ldots, a_n$ des éléments de A non tous nuls et soit $b \in A$. Enfin, soit $d \in A$ le pgcd normalisé de $a_1, \ldots, a_n$ (par rapport à un s.c.r.i. fixé). Alors, l'équation*

$$a_1 x_1 + \cdots + a_n x_n = b, \quad x_i \in A$$

admet une solution si, et seulement si, $d \mid b$. Dans ce cas, toute solution de cette équation s'écrit de manière unique sous la forme

$$\alpha \cdot V_1 + x_2 \cdot V_2 + \cdots + x_n \cdot V_n, \quad x_2, \ldots, x_n \in A,$$

où $\alpha \in A$ vérifie $b = \alpha d$ et $V_1, \ldots, V_n$ sont les colonnes d'une matrice inversible $V \in \mathrm{GL}_n(A)$ vérifiant

$$(a_1 \ \cdots \ a_n)\, V = (d \ 0 \ \cdots \ 0).$$

Démonstration. Soit $C = (a_1 \ \cdots \ a_n)$. Alors, on a $\mu_1(C) = d$. D'après le théorème 2.3, il existe $u \in \mathrm{GL}_1(A) = A^\times$ et $V' \in \mathrm{GL}_n(A)$ telles que

$$uCV' = (d \ 0 \ \cdots \ 0).$$

L'équation considérée possède une solution si, et seulement si, $b \in \mathrm{Im}(C)$. La proposition 4.6, appliquée aux matrices $U = \mathrm{I}_1$ et $V = uV'$, montre que cela équivaut à $b = Ub = d\alpha$ pour un certain $\alpha \in A$, c'est-à-dire $d \mid b$.

Dans ce cas, par la remarque 4.7, une solution particulière est donnée par $X_0 = \alpha V_1$. D'après la proposition 4.8, les $n-1$ dernières colonnes de V forment une base de $\mathrm{Ker}(C)$. Comme l'ensemble des solutions de l'équation est égal à $X_0 + \mathrm{Ker}(C)$, cela achève la démonstration. □

4.10. Exemple. Soit $A = \mathbb{Z}$, et considérons l'équation

$$3x + 4y + 7z = b.$$

Puisque $3, 4$ et 7 sont premiers dans leur ensemble, cette équation a une solution pour tout $b \in \mathbb{Z}$. Nous allons appliquer l'algorithme de réduction pour trouver une matrice $V \in \mathrm{GL}_3(\mathbb{Z})$ telle que

$$(3 \ 4 \ 7)\, V = (1 \ 0 \ 0).$$

Afin de calculer V, nous allons utiliser une matrice augmentée.

Considérons donc la matrice

$$N = \begin{pmatrix} 3 & 4 & 7 \\ x & y & z \end{pmatrix}.$$

On effectue les opérations $C_2 \longleftarrow C_2 - C_1$ et $C_3 \longleftarrow C_3 - 2C_1$ pour obtenir

$$N \sim \begin{pmatrix} 3 & 1 & 1 \\ x & -x+y & -2x+z \end{pmatrix}.$$

On échange les deux premières colonnes, et l'on a

$$N \sim \begin{pmatrix} 1 & 3 & 1 \\ -x+y & x & -2x+z \end{pmatrix}.$$

Enfin, on effectue les opérations $C_2 \longleftarrow C_2 - 3C_1$ et $C_3 \longleftarrow C_3 - C_1$ pour obtenir

$$N \sim \begin{pmatrix} 1 & 0 & 0 \\ -x+y & 4x-3y & -x-y+z \end{pmatrix}.$$

Par construction, la deuxième ligne de la matrice de droite n'est rien d'autre que le vecteur $(x \;\; y \;\; z)\,V$. On obtient ainsi

$$V = \begin{pmatrix} -1 & 4 & -1 \\ 1 & -3 & -1 \\ 0 & 0 & 1 \end{pmatrix},$$

et les solutions de cette équation sont donc

$$\begin{pmatrix} -b+4z_1-z_2 \\ b-3z_1-z_2 \\ z_2 \end{pmatrix}, \quad \text{avec } z_1,\, z_2 \in \mathbb{Z}.$$

Le résultat suivant donne une variante pour trouver une A-base de $\mathrm{Ker}(C)$. Notons tout d'abord qu'il est facile de trouver une famille de $n-r$ vecteurs $X_1, \ldots, X_{n-r}$ appartenant à $\mathrm{Ker}(C)$ qui soit libre : on résout le système $CX = 0$ dans K_A^n, on choisit une base de solutions, et l'on multiplie éventuellement par un élément convenable de A pour s'assurer que les vecteurs de cette base sont dans A^n.

Cela ne fournit hélas pas une base de $\mathrm{Ker}(C)$, comme le montre l'exemple de la matrice $C = 0 \in \mathrm{M}_2(\mathbb{Z})$ et des vecteurs $X_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ et $X_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$. En effet, dans ce cas, (X_1, X_2) est clairement une $\mathbb{Q}$-base de solution du système $CX = 0$ dans $\mathbb{Q}^2$, mais ce n'est pas une $\mathbb{Z}$-base de $\mathrm{Ker}(C) = \mathbb{Z}^2$, puisque la matrice $\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$ n'est pas un élément de $\mathrm{GL}_2(\mathbb{Z})$.

Néanmoins, on peut construire une base de $\mathrm{Ker}(C)$ à partir d'une famille de vecteurs $X_1, \ldots, X_{n-r}$ ayant les propriétés ci-dessus.

4.11. Proposition. *Soit A un anneau euclidien, et soit $C \in \mathrm{M}_{m\times n}(A)$ une matrice de rang r. Enfin, soit $(X_1, \ldots, X_{n-r})$ une famille libre d'éléments de $\mathrm{Ker}(C)$, et soit $N \in \mathrm{M}_{n\times n-r}(A)$ la matrice de rang $n-r$ dont les colonnes sont $X_1, \ldots, X_{n-r}$. Si l'on a*

$$UNV = E(b_1, \ldots, b_{n-r}), \quad \textit{avec } U \in \mathrm{GL}_n(A),\ V \in \mathrm{GL}_{n-r}(A),$$

alors une base de $\mathrm{Ker}(C)$ *est donnée par les* $n-r$ *premières colonnes de la matrice* U^{-1}.

Démonstration. Soient $Z_1, \ldots, Z_n$ les colonnes de U^{-1}. Remarquons que ces vecteurs forment une base de A^n, puisque U est inversible. D'après la proposition 4.6, $b_1 \cdot Z_1, \ldots, b_{n-r} \cdot Z_{n-r}$ est une base de $\mathrm{Im}(N)$. Remarquons également que par construction $\mathrm{Im}(N) \subset \mathrm{Ker}(C)$. On a donc

$$b_i \cdot CZ_i = C(b_i \cdot Z_i) = 0, \quad \text{pour tout } i = 1, \ldots, n-r.$$

Puisque b_i est non nul et que A est intègre, on obtient

$$Z_1, \ldots, Z_{n-r} \in \mathrm{Ker}(C).$$

Ainsi, $\mathrm{Ker}(C)$ contient le sous-module de A^n engendré par $Z_1, \ldots, Z_{n-r}$.

Réciproquement, soit $X \in \mathrm{Ker}(C)$. Puisque $(Z_1, \ldots, Z_n)$ est une base de A^n, on peut écrire $X = \sum_{j=1}^{n} \lambda_j \cdot Z_j$, $\lambda_j \in A$. D'après ce qui précède, on a

$$X' = X - \sum_{j=1}^{n-r} \lambda_j \cdot Z_j \in \mathrm{Ker}(C).$$

Montrons que $X' = 0$. Dans le cas contraire, il est facile de voir que la famille $Z_1, \ldots, Z_{n-r}, X'$ est K_A-libre, ce qui implique que le rang du module $\mathrm{Ker}(C)$ est strictement supérieur à $n-r$, d'où une contradiction avec la proposition 4.8. Ainsi, $X' = 0$ et X est combinaison linéaire de $Z_1, \ldots, Z_{n-r}$. On en déduit donc que $Z_1, \ldots, Z_{n-r}$ engendrent $\mathrm{Ker}(C)$. Comme cette famille est aussi libre, on en déduit le résultat voulu. □

4.12. Exemple. Soit A un anneau euclidien, et soient $a_1, \ldots, a_n \in A$ des éléments non tous nuls. Pour trouver les solutions de l'équation

$$a_1 x_1 + \cdots + a_n x_n = 0,$$

il suffit d'appliquer la proposition précédente à la matrice N dont les vecteurs colonnes sont

$$\begin{pmatrix} a_2 \\ -a_1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \begin{pmatrix} a_3 \\ 0 \\ -a_1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \cdots, \quad \begin{pmatrix} a_n \\ 0 \\ 0 \\ \vdots \\ 0 \\ -a_1 \end{pmatrix}.$$

Considérons à nouveau l'équation $3x + 4y + 7z = 0$. Dans ce cas, on a

$$N = \begin{pmatrix} 4 & 7 \\ -3 & 0 \\ 0 & -3 \end{pmatrix}.$$

Puisque l'on veut au final l'inverse de la matrice U, on a intérêt à

(1) faire au maximum des opérations du type $C_j \longleftarrow C_j + aC_i$ ou des échanges de colonnes, puisque cela consiste à multiplier C à droite par des matrices élémentaires, et donc ne modifie pas U ;

(2) calculer la matrice $U' = U^{-1}$ de proche en proche, chose d'autant plus aisée que les inverses des matrices élémentaires sont très faciles à trouver.

Initialisons par $U'^{(0)} = \mathrm{I}_3$ et $N^{(0)} = N$. L'opération $C_2 \longleftarrow C_2 - 2C_1$ donne

$$N^{(0)} \sim N^{(1)} = \begin{pmatrix} 4 & -1 \\ -3 & 6 \\ 0 & -3 \end{pmatrix}.$$

En échangeant les colonnes, on obtient

$$N^{(1)} \sim N^{(2)} = \begin{pmatrix} -1 & 4 \\ 6 & -3 \\ -3 & 0 \end{pmatrix}.$$

L'opération $C_2 \longleftarrow C_2 + 4C_1$ donne alors

$$N^{(2)} \sim N^{(3)} = \begin{pmatrix} -1 & 0 \\ 6 & 21 \\ -3 & -12 \end{pmatrix}.$$

Puisque l'on a effectué des opérations sur les colonnes, on a

$$U'^{(1)} = U'^{(2)} = U'^{(3)} = \mathrm{I}_3.$$

On échange les deux dernières lignes. On obtient alors

$$N^{(3)} \sim N^{(4)} = \begin{pmatrix} -1 & 0 \\ -3 & -12 \\ 6 & 21 \end{pmatrix}.$$

On a donc $U^{(4)} = S_{2,3}^{(3)} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ et par conséquent

$$U'^{(4)} = \left(S_{2,3}^{(3)}\right)^{-1} = S_{2,3}^{(3)}.$$

Les opérations $L_2 \longleftarrow L_2 - 3L_1$ et $L_3 \longleftarrow L_3 + 6L_1$ nous donnent

$$N^{(4)} \sim N^{(5)} = \begin{pmatrix} -1 & 0 \\ 0 & -12 \\ 0 & 21 \end{pmatrix}.$$

Les matrices de transvections correspondantes sont

$$T_{2,1}^{(3)}(-3) = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{et} \quad T_{3,1}^{(3)}(6) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 6 & 0 & 1 \end{pmatrix}.$$

Ainsi, on a $U^{(5)} = T_{2,1}^{(3)}(-3)\, T_{3,1}^{(3)}(6)\, U^{(4)}$, et donc

$$U'^{(5)} = S_{2,3}^{(3)}\, T_{3,1}^{(3)}(-6)\, T_{2,1}^{(3)}(3) = \begin{pmatrix} 1 & 0 & 0 \\ -6 & 0 & 1 \\ 3 & 1 & 0 \end{pmatrix}.$$

On fait $L_3 \longleftarrow L_3 + 2L_2$ et l'on a

$$N^{(5)} \sim N^{(6)} \begin{pmatrix} -1 & 0 \\ 0 & -12 \\ 0 & -3 \end{pmatrix}.$$

On a donc $U^{(6)} = T_{3,2}^{(3)}(2)\, U^{(5)}$ et

$$U'^{(6)} = U'^{(5)}\, T_{3,2}^{(3)}(-2) = \begin{pmatrix} 1 & 0 & 0 \\ -6 & -2 & 1 \\ 3 & 1 & 0 \end{pmatrix}.$$

On échange les deux dernières lignes pour obtenir

$$N^{(6)} \sim N^{(7)} = \begin{pmatrix} -1 & 0 \\ 0 & -3 \\ 0 & -12 \end{pmatrix},$$

ainsi que $U'^{(7)} = U'^{(6)} S_{2,3}^{(3)} = \begin{pmatrix} 1 & 0 & 0 \\ -6 & 1 & -2 \\ 3 & 0 & 1 \end{pmatrix}$.

Enfin, on fait $L_3 \longleftarrow L_3 - 4L_2$ et lon a

$$N^{(7)} \sim N^{(8)} = \begin{pmatrix} -1 & 0 \\ 0 & -3 \\ 0 & 0 \end{pmatrix}.$$

On a donc $U = U^{(8)} = T_{3,2}^{(3)}(-4)\, U^{(7)}$, et ainsi

$$U^{-1} = U'^{(7)}\, T_{3,2}^{(3)}(4) = \begin{pmatrix} 1 & 0 & 0 \\ -6 & -7 & -2 \\ 3 & 4 & 1 \end{pmatrix}.$$

Une base de solutions est donc donnée par

$$e_1 = \begin{pmatrix} 1 \\ -6 \\ 3 \end{pmatrix}, \quad e_2 = \begin{pmatrix} 0 \\ -7 \\ 4 \end{pmatrix}.$$

Remarquons que les solutions particulières

$$\begin{pmatrix} 4 \\ -3 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 7 \\ 0 \\ -3 \end{pmatrix}$$

s'écrivent bien dans cette base. En effet, on a

$$\begin{pmatrix} 4 \\ -3 \\ 0 \end{pmatrix} = 4e_1 - 3e_2 \quad \text{et} \quad \begin{pmatrix} 7 \\ 0 \\ -3 \end{pmatrix} = 7e_1 - 6e_2.$$

Au passage, cela montre que $\mathrm{Im}(N) \subset \mathbb{Z}\cdot e_1 + 3\mathbb{Z}\cdot e_2$, et donc que les éléments du noyau e_1 et e_2 choisis initialement ne forment pas une base de l'ensemble des solutions, puisque $e_2 \notin \mathrm{Im}(N)$.

4.13. Remarque. Soit A un anneau euclidien. Si $C \in \mathrm{M}_{m\times n}(A)$, il n'est en fait pas nécessaire de trouver sa forme normale de Smith pour résoudre le système linéaire associé. En fait, on peut réduire C sous une forme échelonnée particulière, appelée forme normale d'Hermite, en utilisant des opérations élémentaires sur ses lignes. Le lecteur pourra consulter [19] ou [22] avec profit pour plus de renseignements (les algorithmes sont expliqués dans le cas où $A = \mathbb{Z}$, mais s'adaptent aisément au cas d'un anneau euclidien).

On finit ce chapitre en donnant une nouvelle démonstration de la partie existence du théorème de structure des groupes abéliens de type fini.

4.14. Théorème. *Soit G un groupe abélien de type fini. Alors, il existe un entier $r \geqslant 0$ et des entiers $d_1, \ldots, d_s \geqslant 2$ vérifiant $d_s \mid \cdots \mid d_2 \mid d_1$ tels que*

$$G \simeq \mathbb{Z}/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_s\mathbb{Z} \times \mathbb{Z}^r.$$

De plus, l'entier r et la suite d'entiers $(d_1, \ldots, d_s)$ est unique, et ne dépend que de la classe d'isomorphisme de G.

Démonstration. Soient G un groupe abélien de type fini, et $g_1, \ldots, g_n \in G$ une famille de générateurs de G. On peut toujours supposer que $n \geqslant 1$, quitte à rajouter un élément nul.

Considérons le morphisme de groupes

$$\begin{array}{rccc} f: & \mathbb{Z}^n & \longrightarrow & G \\ & \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} & \longmapsto & \displaystyle\sum_{k=1}^{n} a_k \cdot g_k \cdot \end{array}$$

Par définition, f est un morphisme de groupes surjectif, et le théorème de factorisation donne alors un isomorphisme

$$\mathbb{Z}^n/\mathrm{Ker}(f) \simeq G.$$

Or, $\mathrm{Ker}(f)$ est un sous-groupe de $\mathbb{Z}^n$, donc un sous-module du $\mathbb{Z}$-module libre $\mathbb{Z}^n$. L'anneau $\mathbb{Z}$ étant euclidien, le théorème de la base adaptée nous fournit l'existence d'une $\mathbb{Z}$-base $(e_1, \ldots, e_n)$ de $\mathbb{Z}^n$, d'un entier $t \in [\![0, n]\!]$, et d'entiers $q_1, \ldots, q_t \geqslant 1$ vérifiant $q_1 \mid \cdots \mid q_t$, tels que $(q_1 \cdot e_1, \ldots, q_t \cdot e_t)$ soit une $\mathbb{Z}$-base de $\mathrm{Ker}(f)$.

On vérifie alors sans trop de mal que l'application

$$\begin{array}{rcl} \mathbb{Z}^n & \longrightarrow & \mathbb{Z}/q_1\mathbb{Z} \times \cdots \times \mathbb{Z}/q_t\mathbb{Z} \times \mathbb{Z}^{n-t} \\ \displaystyle\sum_{k=1}^{n} a_k \cdot e_k & \longmapsto & ([a_1]_{q_1}, \ldots, [a_t]_{q_t}, a_{t+1}, \ldots, a_n) \end{array}$$

est surjective, de noyau $\mathrm{Ker}(f)$.

On obtient ainsi des isomorphismes de groupes

$$G \simeq \mathbb{Z}^n/\mathrm{Ker}(f) \simeq \mathbb{Z}/q_1\mathbb{Z} \times \cdots \times \mathbb{Z}/q_t\mathbb{Z} \times \mathbb{Z}^{n-t},$$

d'où l'existence de l'isomorphisme désiré après une renumérotation adéquate. L'unicité ayant déjà été établie auparavant dans la démonstration du théorème XII-5.9, on a fini. □

5. Exercices

E1. Détailler les exemples 3.2.

E2. Soit A un anneau commutatif, que l'on peut même supposer euclidien, et soit M un A-module. Montrer que

(i) si $(x_1, \cdots, x_n)$ est une famille liée d'éléments du A-module M, cela ne veut pas dire qu'un des x_i est combinaison linéaire des autres éléments de la famille, contrairement au cas des espaces vectoriels ;

(ii) Si M est A-libre de rang n, une famille libre de cardinal n n'est pas nécessairement une base ;

(iii) on ne peut pas non plus en général compléter une famille d'éléments linéairement indépendants d'un module libre en une base ;

(iv) on ne peut pas non plus extraire en général une base d'une famille génératrice.

E3. Soit A un anneau possédant des éléments x, y, α, β satisfaisant aux conditions suivantes :

(i) $\alpha x + \beta y = 1$;

(ii) $x\alpha = 1$ et $x\beta = 0$;

(iii) $y\alpha = 0$ et $y\beta = 1$.

a. Montrer que pour tout $r \geqslant 1$, la famille $(y, yx, \ldots, yx^{r-1}, x^r)$ est libre.

b. Montrer que pour tout $r \geqslant 1$, on a

$$1 = \sum_{i=0}^{r-1} (\alpha^i \beta) \cdot y x^i + \alpha^r \cdot x^r.$$

c. En déduire pour tout entier $r \geqslant 1$, la famille $(y, yx, \ldots, yx^{r-1}, x^r)$ est une A-base de A.

d. Soit $A = \mathrm{End}(V)$, où V est un espace vectoriel, de dimension *infinie*, possédant une base dénombrable $(e_n)_{n \geqslant 0}$. On pose

$$x : \begin{array}{ccc} V & \longrightarrow & V \\ \sum_{n \geqslant 0} v_n \cdot e_n & \longmapsto & \sum_{n \geqslant 0} v_{2n+1} \cdot e_n, \end{array} \qquad y : \begin{array}{ccc} V & \longrightarrow & V \\ \sum_{n \geqslant 0} v_n \cdot e_n & \longmapsto & \sum_{n \geqslant 0} v_{2n} \cdot e_n, \end{array}$$

$$\alpha : \begin{array}{ccc} V & \longrightarrow & V \\ \sum_{n \geqslant 0} v_n \cdot e_n & \longmapsto & \sum_{n \geqslant 0} v_n \cdot e_{2n+1}, \end{array} \qquad \beta : \begin{array}{ccc} V & \longrightarrow & V \\ \sum_{n \geqslant 0} v_n \cdot e_n & \longmapsto & \sum_{n \geqslant 0} v_n \cdot e_{2n}. \end{array}$$

Vérifier que x, y, α et β sont des endomorphismes de V satisfaisant les conditions (i)-(iii).

E4. Soit M un A-module, et soit N un sous-module de M.

a. Montrer que l'application

$$\begin{array}{ccc} A \times M/N & \longrightarrow & M/N \\ (a, \overline{x}) & \longmapsto & \overline{a \cdot x} \end{array}$$

est bien définie, et induit sur le groupe abélien $(M/N, +)$ une structure de A-module telle que la projection canonique $\pi : M \longrightarrow M/N$ soit une application A-linéaire.

b. Soit $f : M \longrightarrow M'$ une application A-linéaire telle que $N \subset \mathrm{Ker}(f)$. Montrer qu'il existe une unique application A-linéaire $\overline{f} : M/N \longrightarrow M'$ telle que $f = \overline{f} \circ \pi$.

c. Montrer que pour toute application A-linéaire $f : M \longrightarrow M'$, on a un isomorphisme de A-modules

$$A/\mathrm{Ker}(f) \simeq \mathrm{Im}(f).$$

E5. Soit A un anneau commutatif, et soit M un A-module. On dit que M est *noethérien* si tout sous-module de M est de type fini (en particulier, le A-module M est de type fini).

a. Vérifier qu'un anneau noethérien A est un A-module noethérien (pour la structure de A-module induite par la loi produit).

b. Soit $f : M \longrightarrow N$ une application A-linéaire surjective. Montrer que M est noethérien si, et seulement si, N et $\mathrm{Ker}(f)$ sont noethériens.

Indication. Si N et $\mathrm{Ker}(f)$ sont noethériens, et si M' est un sous-module de M, montrer qu'il existe $x'_1, \ldots, x'_r \in M'$ dont les images par f engendrent $f(M')$. Si $x'_{r+1}, \ldots, x'_s$ engendrent $M' \cap \mathrm{Ker}(f)$, montrer alors que $x'_1, \ldots, x'_s$ engendrent M'. Inversement, si M est noethérien, pour tout sous-module N' de N, $f^{-1}(N')$ est un sous-module de M. Que vaut alors $f\big(f^{-1}(N')\big)$?

c. En déduire que si A est noethérien, alors A^n est un A-module noethérien pour tout $n \geqslant 1$.

d. Soit A un anneau noethérien. Montrer que tout A-module M de type fini est noethérien.

Indication. Considérer une application A-linéaire surjective

$$f : A^n \longrightarrow M,$$

avec un entier $n \geqslant 1$ bien choisi.

E6. Soit M un A-module de type fini.

a. En utilisant l'exercice 4, montrer qu'il existe $n \geqslant 1$ et un sous-module N de A^n tel que $M \simeq A^n/N$.

b. Soit A un anneau euclidien, soit $\mathscr{P}$ un s.c.r.i. de A, soit $n \geqslant 1$ et soit N un sous-module de A^n de rang m. Montrer qu'il existe des éléments normalisés non nuls $a_1, \ldots, a_m \in A \setminus A^\times$ vérifiant $a_1 \mid a_2 \mid \cdots \mid a_m$ tels que

$$A^n/N \simeq A/(a_1) \times \cdots \times A/(a_m) \times A^{n-m}.$$

c. Soit A un anneau euclidien, soit $\mathscr{P}$ un s.c.r.i. et soit M un A-module de type fini. Montrer qu'il existe un entier $r \geqslant 0$ et des éléments

normalisés non nuls[5] $a_1, \ldots, a_s \in A \setminus A^\times$ vérifiant $a_1 \mid a_2 \mid \cdots \mid a_s$ tels que

$$M \simeq A/(a_1) \times \cdots \times A/(a_s) \times A^r.$$

E7. Soit A un anneau commutatif intègre de caractéristique 0 qui est un groupe abélien de type fini. On suppose connus les résultats de l'exercice 6.

a. Montrer que A est un $\mathbb{Z}$-module libre.

Indication. Utiliser la dernière question de l'exercice 6, et montrer que l'intégrité de A implique que $s = 0$.

b. Soit $\mathfrak{a}$ un idéal de A non nul. Justifier que $\mathfrak{a}$ est un $\mathbb{Z}$-module libre, et que $A/\mathfrak{a}$ est fini si, et seulement si, $\mathfrak{a}$ est de même $\mathbb{Z}$-rang que A.

c. Soit $\mathfrak{a}$ un idéal de A non nul, et soit $a \in \mathfrak{a} \setminus \{0\}$.

(i) Montrer que si A est de rang n comme $\mathbb{Z}$-module libre, il en est de même de l'idéal (a).

(ii) En déduire que $A/(a)$ est fini, puis que $A/\mathfrak{a}$ est fini.

d. Montrer alors que pour tout idéal $\mathfrak{a}$ de A non nul, $A/\mathfrak{a}$ est fini, et $\mathfrak{a}$ est un $\mathbb{Z}$-module libre de même $\mathbb{Z}$-rang que A.

e. Montrer alors que tout idéal premier non nul de A est maximal.

Indication. Que dire d'un anneau fini intègre ?

f. Soit $a \in A \setminus \{0\}$.

On souhaite montrer que $|A/(a)| = |\det(\ell_a)|$, où ℓ_a est l'endomorphisme de multiplication à gauche dans le $\mathbb{Z}$-module libre A.

(i) Justifier l'existence d'une $\mathbb{Z}$-base $(w_1, \ldots, w_n)$ de l'anneau A, et d'entiers $q_1, \ldots, q_n \geqslant 1$ tels que $(q_1 w_1, \ldots, q_1 w_n)$ soit une $\mathbb{Z}$-base de aA.

(ii) Soit P la matrice de $(q_1 w_1, \ldots, q_1 w_n)$ dans la $\mathbb{Z}$-base $(aw_1, \ldots, aw_n)$. Justifier que $\det(P) = \pm 1$.

(iii) Soit M la matrice représentative de ℓ_a dans la base $(w_1, \ldots, w_n)$. Que représente la matrice MP ?

(iv) En déduire que $|\det(\ell_a)| = q_1 \times \cdots \times q_n$, puis conclure.

5. Comme dans le cas des groupes abéliens de type fini, qui correspond au cas $A = \mathbb{Z}$, on peut démontrer l'unicité de r et des $a_1, \ldots, a_s$.

Chapitre XXI

Polynômes irréductibles et applications

Dans ce chapitre, nous allons donner quelques critères d'irréductibilité de polynômes à coefficients dans un anneau. Nous nous intéresserons également à une famille de polynômes irréductibles sur $\mathbb{Z}$ et sur $\mathbb{Q}$, à savoir les polynômes cyclotomiques, ainsi qu'à l'existence de polynômes irréductibles sur $\mathbb{F}_p$ de degré fixé.

Dans tout ce qui suit, sauf mention expresse du contraire, A sera un anneau commutatif.

1. Polynômes irréductibles et racines

Par définition, un polynôme $f \in A[X]$ est irréductible s'il est non nul, non inversible, et si pour tous g, $h \in A[X]$, on a

$$f = gh \quad \Longrightarrow \quad g \in A[X]^\times \text{ ou } \ h \in A[X]^\times.$$

Lorsque A est un corps K, d'après la remarque XVIII-1.16, $f \in K[X]$ est irréductible si, et seulement si, il n'est pas constant et ne peut pas s'écrire comme produit de deux polynômes non constants, ce qui est la définition classique de l'irréductibilité d'un polynôme.

Attention ! Cela est faux si A est n'est pas un corps. Par exemple, des considérations de degrés montrent que le polynôme $2X \in \mathbb{Z}[X]$ ne peut s'écrire comme produit de deux polynômes non constants. En revanche, il n'est pas irréductible, puisque 2 et X ne sont pas inversibles dans $\mathbb{Z}[X]$. En effet, d'après le lemme XV-6.31 (2), l'anneau $\mathbb{Z}$ étant intègre, on a

$$\mathbb{Z}[X]^\times = \mathbb{Z}^\times = \{\pm 1\}.$$

Intéressons-nous maintenant à l'irréductibilité des polynômes de petit degré. On peut déjà régler le cas des polynômes constants. Dans ce cas, on a le résultat général suivant.

1.1. Lemme. *Soit A un anneau intègre, et soit $\pi \in A$. Alors, π est irréductible dans $A[X]$ si, et seulement si, π est irréductible dans A.*

Démonstration. On rappelle que, puisque A est intègre, $A[X]$ est aussi intègre et $A[X]^\times = A^\times$. On en déduit facilement que $\pi \in A$ est non nul et non inversible dans A si, et seulement si, il est non nul et non inversible dans $A[X]$.

Supposons que π soit un élément irréductible de A. Puisque π est non nul, non inversible dans A, il est non nul, et non inversible dans $A[X]$. Supposons maintenant que $\pi = fg$, $fg \in A[X]$. En comparant les degrés, on a

$$0 = \deg(\pi) = \deg(fg) = \deg(f) + \deg(g),$$

car A est intègre. On en déduit que f et g sont constants, c'est-à-dire $f = a$ et $g = b$, $a, b \in A$. Comme $\pi = ab$, $a, b \in A$ et que π est irréductible dans A, on a $a \in A^\times$ ou $b \in A^\times$. Ainsi, $f \in A[X]^\times$ ou $g \in A[X]^\times$.

Inversement, supposons que $\pi \in A$ soit irréductible dans $A[X]$. Alors, π est non nul et non inversible dans A. Si maintenant $\pi = ab$, $a, b \in A$, alors en voyant cette égalité dans $A[X]$, on en déduit que $a \in A[X]^\times$ ou $b \in A[X]^\times$, c'est-à-dire $a \in A^\times$ ou $b \in A^\times$. Ainsi, π est irréductible dans A. Cela achève la démonstration. □

Dans le cas des polynômes non constants, l'existence ou non de racines permet parfois de démontrer l'irréductibilité d'un polynôme.

1.2. Exemples

(1) Supposons A **intègre**. Si $f \in A[X]$ possède une racine dans A, et si l'on a $\deg(f) \geqslant 2$, alors f n'est pas irréductible.
En effet, si $f(a) = 0$ avec $a \in A$, alors $f = (X - a)g$, $g \in A[X]$. Or, $X - a$ et g ne sont pas constants (pour des questions de degré), donc sont non inversibles, puisque $A[X]^\times = A^\times$. Ainsi, f n'est pas irréductible.
C'est faux si A n'est pas intègre. En effet, $X \in \mathbb{Z}/4\mathbb{Z}[X]$ est irréductible (ce n'est pas totalement immédiat. On peut par exemple utiliser le théorème 1.7). Comme $\overline{2}X + \overline{1}$ est inversible (car $(\overline{2}X + \overline{1})^2 = \overline{1}$), alors $f = (\overline{2}X + \overline{1})X$ est irréductible de degré 2, mais il possède une racine.

(2) Si K est un corps, et si $f \in K[X]$ est de degré 2 ou 3, alors f est irréductible si, et seulement si, f n'a pas de racines dans K.

En effet, si f admet un facteur non constant, il en aura nécessairement un de degré 1.

Cela est faux si $\deg(f) \geqslant 4$, puisque $(X^2+1)^2 \in \mathbb{R}[X]$ n'a pas de racines, mais n'est pas irréductible.

Cela est aussi faux si K n'est pas un corps, puisque $2(X^2+1) \in \mathbb{Z}[X]$ n'a pas de racines, mais n'est pas irréductible.

Lorsque K est un corps, tout polynôme de degré 1 à coefficients dans K est irréductible. Plus généralement, on a le résultat suivant.

1.3. Lemme. *Soit A un anneau intègre. Alors, pour tout $a \in A$, le polynôme $X - a$ est irréductible.*

Démonstration. Soit $a \in A$. Alors, $X - a$ est non nul, et non inversible, car A étant intègre, on a $A[X]^\times = A^\times$. Supposons que l'on puisse écrire

$$X - a = fg, \quad f, g \in A[X].$$

Remarquons que f et g sont non nuls. Comme A est intègre, on a

$$\deg(X - a) = 1 = \deg(f) + \deg(g).$$

En particulier, f ou g est constant, disons f. Dans ce cas, g est de degré 1. Écrivons $f = b$ et $g = cX + d$. En identifiant les coefficients dominants, on obtient $bc = 1$. Ainsi, $f = b \in A^\times$ et $X - a$ est bien irréductible. □

1.4. Remarque. On a en fait un peux mieux. Si A est intègre, alors $X - a$ est un élément premier de $A[X]$ pour tout $a \in A$. En effet, le morphisme d'évaluation en a induit un isomorphisme d'anneaux

$$A[X]/(X - a) \simeq A,$$

et $(X - a)$ est un idéal premier. En particulier, on retrouve le fait que $X - a$ est irréductible par le lemme XVIII-1.19.

L'exemple suivant montre que ce résultat n'est plus valable si A n'est plus intègre.

1.5. Exemple. Soit $A = \mathbb{Z}/6\mathbb{Z}$. Alors, $\overline{3}X - \overline{2}$ et $-\overline{2}X + \overline{3}$ ne sont pas inversibles dans $A[X]$. En effet, si $f \in A[X]^\times$, alors $f(0) \in A^\times$, par la proposition [1] XV-6.37. Néanmoins, on a

$$X = (\overline{3}X - \overline{2})(-\overline{2}X + \overline{3}).$$

Ainsi, X n'est pas irréductible.

1. On peut aussi le démontrer directement : si $fg = 1 \in A[X]$, alors $f(0)g(0) = 1 \in A$.

Nous allons maintenant caractériser les anneaux A pour lesquels X est irréductible. Auparavant, signalons le critère d'irréductibilité suivant, qui est une conséquence immédiate du lemme XVIII-1.24, appliqué à l'isomorphisme d'anneaux

$$\begin{aligned} \tau_a : A[X] &\longrightarrow A[X] \\ f &\longmapsto f(X-a). \end{aligned}$$

1.6. Lemme. *Soit $f \in A[X]$ et soit $a \in A$. Alors, f est irréductible si, et seulement si, $f(X-a)$ est irréductible.*

On peut maintenant énoncer le résultat suivant.

1.7. Théorème. *Les propriétés suivantes sont équivalentes :*

(1) *pour tout $a \in A$, $X-a$ n'est pas irréductible ;*

(2) *il existe $a \in A$ tel que $X-a$ ne soit pas irréductible ;*

(3) *le polynôme X n'est pas irréductible ;*

(4) *l'anneau A admet un idempotent e non trivial ;*

(5) *il existe deux anneaux A_1, A_2 non triviaux tels que $A \simeq A_1 \times A_2$.*

Démonstration. L'équivalence (1) $\iff$ (3) découle du lemme 1.6, de même que l'implication (2) $\implies$ (3). De plus, l'implication (3) $\implies$ (2) est évidente. Enfin, l'équivalence (4) $\iff$ (5) est tout simplement l'énoncé du corollaire XV-5.7 dans le cas où A est commutatif.

Il reste donc à démontrer (3) $\iff$ (4). Remarquons tout d'abord que X est non nul, et non inversible (car son terme constant est non inversible). Supposons maintenant que X ne soit pas irréductible. Puisque X est non nul et non inversible, cela implique que l'on peut écrire $X = RS$, où R et S sont non inversibles.

Écrivons $R = a+XP, S = b+XQ$. On a donc $X = ab+(aQ+bP)X+X^2PQ$.

L'évaluation en 0 nous donne $ab = 0$. On a donc $X = (aQ+bP)X+X^2PQ$. Puisque la multiplication par X est injective, on obtient

$$1 = aQ + bP + XPQ.$$

Posons $v = P(0)$ et $u = Q(0)$. En évaluant en 0, on obtient $au + bv = 1$. Soit alors $e = au$. On a

$$e^2 = au(1-bv) = au - abuv = au = e,$$

et e est donc un idempotent. Supposons que $e = 0$. Alors, $bv = 1$, donc b est inversible et ainsi $a = 0$.

On a donc $X = XPS$. Comme la multiplication par X est injective, on obtient $1 = PS$, donc S est inversible, d'où une contradiction. Par conséquent, $e \neq 0$. De la même manière, si $e = 1$, alors $au = 1$ et l'on obtient cette fois que R est inversible.

Ainsi, e est un idempotent non trivial.

Inversement, si e est un idempotent non trivial de A, $eX+1-e$ et $(1-e)X+e$ ne sont pas inversibles, car leurs coefficients constants sont non inversibles (en effet, si e ou $(1-e)$ était inversible, l'égalité $e(1-e) = 0$ entraînerait que $e = 0$ ou 1), mais on a

$$X = (eX + 1 - e)\big((1-e)X + e\big).$$

Ainsi, X n'est pas irréductible et cela achève la démonstration. □

Pour finir ce paragraphe, revenons à des considérations sur les racines de polynômes. Comme on l'a déjà constaté, un polynôme irréductible $f \in K[X]$ n'a pas de racines dans K. Or, l'expérience montre qu'il est parfois nécessaire[2] d'introduire un nouvel élément α vérifiant $f(\alpha) = 0$. Bien évidemment, un tel élément ne peut appartenir au corps K, et la difficulté est de construire un nouveau corps L contenant K (afin de pouvoir considérer f comme un polynôme à coefficients dans L) et dans lequel f possède au moins une racine.

Par exemple, considérons le cas du polynôme $X^2 + 1 \in \mathbb{R}[X]$. Dans ce cas, un exemple de corps L satisfaisant les propriétés précédentes est le corps des nombres complexes. On peut déjà en recenser deux constructions ad hoc. La première façon de construire un corps « contenant $\mathbb{R}$ » et un élément de carré -1 est de définir $\mathbb{C}$ comme le sous-espace vectoriel réel de $\mathrm{M}_2(\mathbb{R})$

$$\mathbb{C} = \left\{ \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \;\middle|\; a, b \in \mathbb{R} \right\}.$$

C'est un espace vectoriel réel de dimension 2, de base

$$1_{\mathbb{C}} = \mathrm{I}_2, \quad i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}.$$

Ainsi, tout élément de $\mathbb{C}$ s'écrit de manière unique

$$a \cdot 1_{\mathbb{C}} + b \cdot i, \quad a, b \in \mathbb{R}.$$

On vérifie que $i^2 = -1_{\mathbb{C}}$ et que $\mathbb{C}$ est un corps, et une $\mathbb{R}$-algèbre. La droite réelle engendrée par $1_{\mathbb{C}}$ est alors un sous-corps de $\mathbb{C}$ isomorphe à $\mathbb{R}$.

2. Voire extrêmement utile.

L'autre construction, un peu plus mystérieuse, est de considérer le groupe abélien $(\mathbb{R}^2, +)$ et de définir une loi produit en posant

$$(a,b)(a',b') = (aa' - bb', ab' + ba'), \quad \text{pour tous } a,\, a',\, b,\, b' \in \mathbb{R}.$$

On vérifie que l'on obtient une $\mathbb{R}$-algèbre qui est aussi un corps, noté $\mathbb{C}$, et vérifiant les propriétés demandées. Dans cette construction, on a $1_{\mathbb{C}} = (1,1)$ et l'élément $i = (0,1)$ vérifie $i^2 = -1_{\mathbb{C}}$.

Ces deux $\mathbb{R}$-algèbres sont d'ailleurs isomorphes, l'isomorphisme faisant correspondre (a,b) et $\begin{pmatrix} a & -b \\ b & a \end{pmatrix}$.

Nous allons maintenant utiliser le corollaire XVIII-2.8 pour détailler une construction plus conceptuelle de $\mathbb{C}$. On aimerait formellement ajouter au corps $\mathbb{R}$ un élément i de carré -1. Autrement dit, on aimerait forcer la relation $i^2 + 1 = 0$. Par conséquent, un tel élément doit être une racine du polynôme $X^2 + 1 \in \mathbb{R}[X]$. Comme dans le cas des groupes, l'idée naturelle est de considérer un quotient approprié.

Posons $\mathbb{C} = \mathbb{R}[X]/(X^2+1)$. D'après le corollaire XVIII-2.8, $\mathbb{C}$ est un corps, puisque $X^2 + 1 \in \mathbb{R}[X]$ est irréductible. Posons $i = \overline{X}$. On a alors

$$i^2 + 1_{\mathbb{C}} = \overline{X}^2 + \overline{1} = \overline{X^2 + 1} = \overline{0} = 0_{\mathbb{C}}.$$

On vérifie ensuite que la loi

$$\begin{aligned} \mathbb{R} \times \mathbb{C} &\longrightarrow \quad \mathbb{C} \\ (\lambda, \overline{P}) &\longmapsto \lambda \cdot \overline{P} = \overline{\lambda P} \end{aligned}$$

induit sur $\mathbb{C}$ une structure de $\mathbb{R}$-espace vectoriel.

Remarquons que $\overline{X^m} = \overline{X}^m = \pm 1_{\mathbb{C}}$ ou $\pm i$ pour tout $m \geqslant 0$. On en déduit que pour tout $P \in \mathbb{R}[X]$, il existe $a,\, b \in \mathbb{R}$ tel que

$$\overline{P} = \overline{a + bX} = a \cdot 1_{\mathbb{C}} + b \cdot i.$$

Ainsi, la famille $(1, i)$ est génératrice. De plus, on a

$$a \cdot 1_{\mathbb{C}} + b \cdot i = 0_{\mathbb{C}} \iff \overline{a + bX} = \overline{0} \iff X^2 + 1 \mid a + bX.$$

Cela équivaut à $a = b = 0$, et $(1, i)$ est donc une $\mathbb{R}$-base de $\mathbb{C}$.

Ainsi, $\mathbb{C}$ est une $\mathbb{R}$-algèbre de dimension 2, ainsi qu'un corps contenant un élément i de carré $-1_{\mathbb{C}}$. Il contient aussi un sous-corps isomorphe à $\mathbb{R}$, à savoir $\mathbb{R} \cdot 1_{\mathbb{C}}$.

Le lecteur vérifiera que les trois couples $(\mathbb{C}, i)$ construits précédemment jouïssent de la propriété suivante.

(CU) Pour toute $\mathbb{R}$-algèbre $\mathscr{A}$ associative, commutative, unitaire, contenant un élément $a \in \mathscr{A}$ vérifiant $a^2 = -1_{\mathscr{A}}$, il existe un unique morphisme de $\mathbb{R}$-algèbres $f : \mathbb{C} \longrightarrow \mathscr{A}$ tel que $f(i) = a$.

Le lemme suivant explique alors pourquoi toutes ces constructions sont isomorphes.

1.8. Lemme. *Soit $\mathbb{C}$ une $\mathbb{R}$-algèbre associative, commutative, unitaire, contenant un élément i tel que $i^2 = -1_{\mathbb{C}}$. Si $(\mathbb{C}, i)$ a la propriété (CU), il existe un unique isomorphisme de $\mathbb{R}$-algèbres $f : \mathbb{R}[X]/(X^2+1) \longrightarrow \mathbb{C}$ tel que $f(\overline{X}) = i$. En particulier, $\mathbb{C}$ est un corps.*

Démonstration. Puisque $(\mathbb{R}[X]/(X^2+1), \overline{X})$ et $(\mathbb{C}, i)$ vérifient la propriété (CU), il existe deux morphismes de $\mathbb{R}$-algèbres $f : \mathbb{R}[X]/(X^2+1) \longrightarrow \mathbb{C}$ et $g : \mathbb{C} \longrightarrow \mathbb{R}[X]/(X^2+1)$ tels que $f(\overline{X}) = i$ et $g(i) = \overline{X}$. Or, les deux morphismes de $\mathbb{R}$-algèbres $f \circ g$ et $\mathrm{Id}_{\mathbb{C}}$ envoient i sur i. Puisque $(\mathbb{C}, i)$ vérifie (CU), on obtient $f \circ g = \mathrm{Id}_{\mathbb{C}}$. De même, on a $g \circ f = \mathrm{Id}_{\mathbb{R}[X]/(X^2+1)}$. Ainsi, f est un isomorphisme de $\mathbb{R}$-algèbres. Son unicité découlant de la propriété (CU), cela achève la démonstration. □

Nous allons maintenant généraliser un peu ces considérations. Soit K un corps, et soit $f \in K[X]$ un polynôme de degré $d \geqslant 1$. La loi externe

$$\begin{aligned} K \times K[X]/(f) &\longrightarrow K[X]/(f) \\ (\lambda, \overline{P}) &\longmapsto \lambda \cdot \overline{P} = \overline{\lambda P} \end{aligned}$$

induit sur $K[X]/(f)$ une structure de K-espace vectoriel. On vérifie aisément que le produit dans l'anneau $K[X]/(f)$ est K-bilinéaire, et donc que l'anneau quotient $K[X]/(f)$ est une K-algèbre (associative unitaire).

On a alors le résultat crucial suivant.

1.9. Proposition. *Soit K un corps, et soit $f \in K[X]$ un polynôme de degré $d \geqslant 1$. Alors, l'anneau quotient $K[X]/(f)$ est une K-algèbre de dimension d, de base $(\overline{1}, \overline{X}, \ldots, \overline{X^{d-1}})$.*

De plus, $K[X]/(f)$ est intègre si, et seulement si, f est irréductible. Dans ce cas, c'est un corps.

Démonstration. Soit $\overline{P} \in K[X]/(f)$. Effectuons la division euclidienne de P par f. Alors, $P = fg + R$, $\deg(R) \leqslant d-1$. On a donc $\overline{P} = \overline{R}$. Si l'on écrit maintenant $R = a_{d-1}X^{d-1} + \cdots + a_1 X + a_0$, on obtient

$$\overline{P} = \overline{R} = \overline{a_{d-1}X^{d-1} + \cdots + a_1 X + a_0} = a_{d-1} \cdot \overline{X^{d-1}} + \cdots + a_1 \cdot \overline{X} + a_0 \cdot \overline{1}.$$

La famille $(\overline{1}, \overline{X}, \ldots, \overline{X^{d-1}})$ est donc génératrice. Pour montrer qu'elle est

libre, soient $a_0, \ldots, a_{d-1} \in K$ tels que

$$a_{d-1} \cdot \overline{X^{d-1}} + \cdots + a_1 \cdot \overline{X} + a_0 \cdot \overline{1} = \overline{0}.$$

Cela se récrit $\overline{R} = \overline{0}$, où $R = a_{d-1}X^{d-1} + \cdots + a_1 X + a_0$. Ainsi, $f \mid R$. Mais, on a $\deg(R) \leqslant d-1$, et donc cela ne peut se produire que si $R = 0$, c'est-à-dire si $a_0 = a_1 = \cdots = a_{d-1} = 0$. Ainsi, la famille $(\overline{1}, \overline{X}, \ldots, \overline{X^{d-1}})$ est libre, et c'est donc une base de $K[X]/(f)$, qui est alors de dimension d.

Le reste découle du corollaire XVIII-2.8. □

Le lemme suivant montre que le corps $K[X]/(f)$ a les propriétés attendues.

1.10. Lemme. *Soit K un corps, et soit $f = a_n X^n + \cdots + a_1 X + a_0$ un polynôme irréductible de $K[X]$. Alors, $K[X]/(f)$ est un corps dans lequel le polynôme $\overline{a}_n X^n + \cdots + \overline{a}_1 X + \overline{a}_0$ a une racine.*

Démonstration. Puisque f est irréductible, $K[X]/(f)$ est un corps par la proposition 1.9. Posons $\alpha = \overline{X} \in K[X]/(f)$. Alors, on a

$$\begin{aligned} \overline{a}_n \alpha^n + \cdots + \overline{a}_1 \alpha + \overline{a}_0 &= \overline{a}_n \overline{X}^n + \cdots + \overline{a}_1 \overline{X} + \overline{a}_0 \\ &= \overline{a}_n \overline{X^n} + \cdots + \overline{a}_1 \overline{X} + \overline{a}_0 \\ &= \overline{g} \\ &= \overline{0}. \end{aligned}$$

Cela achève la démonstration. □

1.11. Remarque. En fait, nous avons un peu exagéré ici. Le corps $K[X]/(f)$ ne contient pas K stricto sensu, mais un sous-corps isomorphe à K, qui est l'image de K par l'application $\iota : K \longrightarrow K[X]/(f)$. Enfin, ce n'est pas vraiment le polynôme f qui possède une racine dans $K[X]/(f)$, mais le polynôme $\overline{a}_n X^n + \cdots + \overline{a}_1 X + \overline{a}_0$.

Il faut travailler un tout petit peu plus afin d'obtenir un corps ayant vraiment les propriétés demandées. Nous reviendrons sur ce point dans un chapitre ultérieur.

2. Quelques critères d'irréductibilité

Dans ce paragraphe, nous allons énoncer des critères d'irréductibilité valables lorsque A est un anneau intègre arbitraire. On commence par introduire la notion de polynôme primitif.

2.1. Définition. On dit que $f = a_n X^n + \cdots + a_1 X + a_0 \in A[X]$ est *primitif* si les coefficients de f sont premiers entre eux dans leur ensemble, i.e. si pour tout $d \in A$, on a

$$d \mid a_i, \quad \text{pour tout } i \in [\![0, n]\!] \quad \Longrightarrow \quad d \in A^\times.$$

2.2. Exemples

(1) Tout polynôme unitaire de $A[X]$ est primitif.

(2) Le polynôme $2X + 3 \in \mathbb{Z}[X]$ est primitif.

L'introduction de cette notion de polynôme primitif est motivée par la remarque suivante.

2.3. Remarque. Si A est intègre, un polynôme irréductible $f \in A[X]$ non constant est primitif.
En effet, soit $d \in A$ un diviseur commun des coefficients de f. Alors, on peut écrire $f = dg$, où $g \in A[X]$ est non constant. Mais, comme A est intègre, on a $A[X]^\times = A^\times$, et g n'est donc pas inversible. Comme f est irréductible, on a donc $d \in A^\times$, ce qu'il fallait vérifier.

On peut maintenant énoncer un premier critère d'irréductibilité.

2.4. Théorème (Critère d'irréductibilité par réduction modulo un idéal). *Soit A un anneau intègre, et soit $\mathfrak{a}$ un idéal de A. Soit $f \in A[X]$ un polynôme primitif non constant. On suppose que f vérifie les conditions suivantes :*

(1) *la classe d'équivalence du coefficient dominant de f est non nulle et non diviseur de zéro dans $A/\mathfrak{a}$;*

(2) *le polynôme $\overline{f} \in (A/\mathfrak{a})[X]$ est irréductible dans $(A/\mathfrak{a})[X]$.*

Alors, f est irréductible dans $A[X]$.

Démonstration. Puisque f est non constant, f est non nul et non inversible. Supposons que l'on ait $f = gh$, $g, h \in A[X]$, et supposons que g et h ne soient pas inversibles. Si g ou h était constant, comme f est primitif, il serait inversible. On a donc $\deg(g) \geqslant 1$ et $\deg(h) \geqslant 1$.
Écrivons

$$g = b_n X^n + \cdots + b_0, \quad h = c_m X^m + \cdots + c_0, \quad b_n, c_m \in A \setminus \{0\},$$

avec $n, m \geqslant 1$. Notons $a_d \in A \setminus \{0\}$ le coefficient dominant de f. Alors, on a $a_d = b_n c_m$. Comme $\overline{a}_d = \overline{b}_n \overline{c}_m \neq \overline{0}$ dans $A/\mathfrak{a}$, on a $\overline{b}_n \neq \overline{0}$ et $\overline{c}_m \neq \overline{0}$ dans $A/\mathfrak{a}$. En particulier, on a $\deg(\overline{g}) \geqslant 1$, $\deg(\overline{h}) \geqslant 1$ et $\overline{f} = \overline{g}\,\overline{h}$.

Montrons que ni $\overline{g}$, ni $\overline{h}$ ne sont inversibles. Supposons par exemple que $\overline{g}$ soit inversible. Alors, il existe $\overline{g}_0 \in A/\mathfrak{a}[X]$ tel que $\overline{g}\,\overline{g}_0 = \overline{1}$. Si $\overline{\delta} \neq \overline{0}$ est le coefficient dominant de $\overline{g}_0$, on aurait $\overline{b}_n \overline{\delta} = \overline{0}$, car $\deg(\overline{g}) \geqslant 1$. Mais alors, on aurait

$$\overline{a}_d\,\overline{\delta} = \overline{c}_m\,\overline{b}_n\,\overline{\delta} = \overline{0}.$$

Autrement dit, le coefficient dominant de f serait un diviseur de zéro dans $A/\mathfrak{a}$, d'où une contradiction. Ainsi, $\overline{g}$ est non inversible, et de même $\overline{h}$ est non inversible. Cela contredit l'irréductibilité de $\overline{f}$, et achève la démonstration. □

2.5. Remarques

(1) Ce critère est particulièrement pratique dans le cas d'un polynôme unitaire. Il se formule alors comme suit : soit $f \in A[X]$ un polynôme unitaire non constant et soit $\mathfrak{a}$ un idéal de A. Si $\overline{f} \in (A/\mathfrak{a})[X]$ est irréductible, alors f est irréductible.

(2) Le résultat est faux si A n'est plus supposé intègre.
En effet, soit $A = \mathbb{Z}/6\mathbb{Z}$, et soit $f = X \in A[X]$. Alors, f est unitaire, donc primitif, et non constant. Soit $\mathfrak{a} = (\overline{2})$. On a aisément l'isomorphisme d'anneaux

$$A/\mathfrak{a} \simeq \mathbb{Z}/2\mathbb{Z}.$$

La classe du coefficient dominant de f est envoyée sur la classe de 1 modulo $\mathfrak{a}$, qui est non nulle et non diviseur de 0. De plus, $A/\mathfrak{a}$ étant isomorphe à $\mathbb{Z}/2\mathbb{Z}$, c'est un anneau intègre et X est donc irréductible modulo $\mathfrak{a}$. En revanche, A possède un idempotent non trivial, à savoir $\overline{3}$. D'après le lemme 1.6, X n'est pas irréductible dans $A[X]$.

Si $\mathfrak{p}$ est un idéal premier, alors $A/\mathfrak{p}$ est intègre et n'a donc pas de diviseurs de zéro. On obtient alors le cas particulier important suivant.

2.6. Corollaire. *Soit A un anneau intègre, et soit $\mathfrak{p}$ un idéal premier de A. Soit $f \in A[X]$ un polynôme primitif non constant. On suppose que f vérifie les conditions suivantes :*

(1) *le coefficient dominant de f n'est pas dans $\mathfrak{p}$;*

(2) *le polynôme $\overline{f} \in (A/\mathfrak{p})[X]$ est irréductible dans $(A/\mathfrak{p})[X]$.*

Alors, f est irréductible dans $A[X]$.

En pratique, on utilise souvent la version suivante, qui correspond au cas où $\mathfrak{p}$ est principal.

2.7. Corollaire. *Soit A un anneau intègre, et soit π un élément premier de A. Soit $f \in A[X]$ un polynôme primitif non constant. On suppose que f vérifie les conditions suivantes :*

(1) *le coefficient dominant de f n'est pas divisible par π ;*

(2) *le polynôme $\overline{f} \in \big(A/(\pi)\big)[X]$ est irréductible dans $\big(A/(\pi)\big)[X]$.*

Alors, f est irréductible dans $A[X]$.

En particulier, on obtient le résultat suivant.

2.8. Corollaire. *Soit $f \in \mathbb{Z}[X]$ un polynôme primitif non constant et soit p un nombre premier. On suppose que*

(1) *le coefficient dominant de f n'est pas divisible par p ;*

(2) *le polynôme $\overline{f}$ est irréductible dans $\mathbb{F}_p[X]$.*

Alors, f est irréductible dans $\mathbb{Z}[X]$.

2.9. Exemple. Le polynôme $f = X^3 + 9X^2 - 10X + 22 \in \mathbb{Z}[X]$ est irréductible. En effet, il est primitif, sa réduction $\overline{f} = X^3 - X + 1 \in \mathbb{Z}/3\mathbb{Z}[X]$ est de degré 3, et irréductible puisque sans racines dans $\mathbb{Z}/3\mathbb{Z}$.

On continue par le critère d'Eisenstein.

2.10. Théorème (Critère d'irréductibilité d'Eisenstein). *Soit A un anneau intègre, et soit $\mathfrak{p}$ un idéal premier de A.*

Soit $f \in A[X]$, $f = a_n X^n + a_{n-1}X^{n-1} + \cdots + a_0$ un polynôme primitif non constant. On suppose que f vérifie les conditions suivantes :

(1) *on a $a_n \notin \mathfrak{p}$;*

(2) *pour tout $i \in [\![0, n-1]\!]$, on a $a_i \in \mathfrak{p}$;*

(3) *on a $a_0 \notin \mathfrak{p}^2$.*

Alors, f est irréductible dans $A[X]$.

Démonstration. Remarquons que f est non nul et non inversible (car il est non constant). Supposons que f ne soit pas irréductible. Alors, on peut écrire $f = gh$, $g, h \in A[X]$. De plus, g et h sont non constants (on l'a déjà constaté dans la démonstration du critère de réduction modulo un idéal).

Considérons la réduction modulo $\mathfrak{p}$. Soient a_n, b_m et c_ℓ les coefficients dominants de f, g et h respectivement. Alors, on a $a_n = b_m c_\ell$, car A est intègre, et par suite $\overline{a}_n = \overline{b}_m\, \overline{c}_\ell$. Par hypothèse, $\overline{a}_n \neq \overline{0}$, et par conséquent, on a $\overline{b}_m \neq \overline{0}$ et $\overline{c}_\ell \neq \overline{0}$. Ainsi, $\deg(\overline{g}) \geqslant 1$ et $\deg(\overline{h}) \geqslant 1$.

Par hypothèse sur f, on a $\overline{f} = \overline{a}_n X^n \in (A/\mathfrak{p})[X], \overline{a}_n \neq \overline{0}$. Comme l'anneau $A/\mathfrak{p}$ est intègre, on peut considérer son corps des fractions F. En voyant l'égalité $\overline{f} = \overline{g}\,\overline{h}$ précédente dans $F[X]$, on obtient

$$\overline{g} = \lambda X^m, \quad \overline{h} = \lambda' X^l, \quad \lambda, \lambda' \in F^\times, \ m + \ell = n.$$

Écrivons $\lambda = \dfrac{\overline{u}}{\overline{v}}$, $\overline{v} \neq \overline{0}$. On a alors $\overline{v}\,\overline{g} = \overline{u} X^r$. Si b_0 désigne le terme constant de g, on a donc $\overline{v}\,\overline{b}_0 = \overline{0}$, car $m \geqslant 1$, et comme $A/\mathfrak{p}$ est intègre, on en déduit que $\overline{b}_0 = \overline{0}$. Autrement dit, $b_0 \in \mathfrak{p}$. De même, le terme constant c_0 de h est un élément de $\mathfrak{p}$. Mais alors $b_0 c_0 \in \mathfrak{p}^2$, ce qui est une contradiction, car $b_0 c_0$ est le terme constant de f, d'après l'égalité $f = gh$. Cela achève la démonstration. □

2.11. Remarque. Encore une fois, ce critère est faux si A n'est pas intègre. Soit $A = \mathbb{Z}/20\mathbb{Z}$, et soit $f = X - \overline{2} \in A[X]$. Alors, f est primitif non constant, et $\mathfrak{p} = (\overline{2})$ est un idéal premier, puisque $A/\mathfrak{p} \simeq \mathbb{Z}/2\mathbb{Z}$.
Remarquons que $\mathfrak{p}^2 = (\overline{4})$, et que $\overline{2} \notin (\overline{4})$. En effet, dans le cas contraire, il existerait $m \in \mathbb{Z}$ tel que $2 \equiv 4m$ [20], d'où $1 \equiv 2m$ [10]. Ainsi, 2 serait inversible modulo 10, ce qui n'est pas le cas.
De plus, f est irréductible modulo $\mathfrak{p}$, puisque $\overline{f} = X$ et que $A/\mathfrak{p}$ est intègre. En revanche, f n'est pas irréductible d'après le lemme 1.6, car A possède un idempotent non trivial, à savoir $\overline{5}$.

En pratique, on utilise souvent la version suivante, que l'on obtient en appliquant le critère précédent à l'idéal engendré par un élément premier de A.

2.12. Corollaire. *Soit A un anneau intègre, et soit π un élément premier de A. Soit $f \in A[X]$, $f = a_n X^n + \cdots + a_1 X + a_0$ un polynôme primitif non constant. On suppose que f vérifie les conditions suivantes :*
(1) *on a $\pi \nmid a_n$;*
(2) *pour tout $i \in [\![0, n-1]\!]$, on a $\pi \mid a_i$;*
(3) *on a $\pi^2 \nmid a_0$.*
Alors, f est irréductible dans $A[X]$.

2.13. Définition. Un polynôme vérifiant les conditions du corollaire s'appelle un *polynôme d'Eisenstein* en π, ou π-Eisenstein.

2.14. Exemple. L'élément $Y - 1 \in \mathbb{Z}[X, Y] = \mathbb{Z}[X][Y]$ est premier dans l'anneau $\mathbb{Z}[X, Y]$, puisque l'évaluation en 1 induit un isomorphisme

$$\mathbb{Z}[X, Y]/(Y - 1) \simeq \mathbb{Z}[X],$$

qui est intègre. Le polynôme $f = YX^4 + Y - 1 \in \mathbb{Z}[X, Y] = \mathbb{Z}[Y][X]$ est alors un polynôme d'Eisenstein en $Y - 1$, et est donc irréductible.

Lorsque $A = \mathbb{Z}$, le critère d'Eisenstein se retraduit comme suit, puisque un idéal premier non nul est engendré par un nombre premier.

2.15. Corollaire. *Soit $f = a_n X^n + \cdots + a_1 X + a_0 \in \mathbb{Z}[X]$, $a_n \neq 0$, un polynôme primitif non constant et soit p un nombre premier.*
On suppose que
(1) *on a $p \nmid a_n$;*
(2) *pour tout $i \in [\![0, n-1]\!]$, on a $p \mid a_i$;*
(3) *on a $p^2 \nmid a_0$.*
Alors, f est irréductible dans $\mathbb{Z}[X]$.

2.16. Exemple. Pour tout $n \geqslant 1$, $X^n - 2 \in \mathbb{Z}[X]$ est 2-Eisenstein, donc irréductible dans $\mathbb{Z}[X]$.

Comme on le verra dans les chapitres ultérieurs, savoir établir l'irréductibilité d'un polynôme f à coefficients dans un corps K est d'une importance vitale. Lorsque K est le corps des fractions K_A d'un anneau intègre A, on peut toujours se réduire au cas où f est à coefficients dans $A[X]$ et unitaire. En effet, si $f \in K_A[X]$, on peut toujours écrire

$$f = \frac{a_n}{d_n} X^n + \cdots + \frac{a_1}{d_1} X + \frac{a_0}{d_0}, \quad a_i \in A,\ d_i \in A \setminus \{0\},$$

avec a_n non nul. Si $d = d_0 \cdots d_n$, on a alors

$$df = b_n X^n + \cdots + b_1 X + b_0, \quad b_i \in A,$$

avec b_n non nul. Enfin, on a

$$b_n^{n-1} df = g(b_n X),$$

avec $g \in A[X]$ unitaire. Or, f est irréductible dans $K_A[X]$ si, et seulement si, $b_n^{n-1} df$ l'est. Mais, l'application

$$\begin{aligned} K_A[X] &\longrightarrow K_A[X] \\ h &\longmapsto h(b_n X) \end{aligned}$$

est un isomorphisme d'anneaux, car $b_n \neq 0$. Vu le lemme XVIII-1.24, on en déduit in fine que f est irréductible dans $K_A[X]$, si, et seulement si, g l'est.

Il est donc naturel de se poser la question suivante : soit A un anneau intègre, et soit $f \in A[X]$ un polynôme unitaire non constant. Y a-t-il un lien entre l'irréductibilité de f dans $A[X]$ et dans $K_A[X]$?

Par la remarque 2.3, si f est un polynôme irréductible non constant de $A[X]$, il est primitif. On a alors le résultat positif suivant, qui s'applique en particulier aux polynômes unitaires.

2.17. Lemme. *Soit A un anneau intègre, et soit $f \in A[X]$ un polynôme primitif non constant. Si f est irréductible dans $K_A[X]$, alors f est irréductible dans $A[X]$.*

Démonstration. Si f est irréductible dans $K_A[X]$, il est en particulier non constant, donc non nul et non inversible dans $A[X]$, car $A[X]^\times = A^\times$. Si maintenant $f = gh$, avec $g, h \in A[X]$, alors l'hypothèse sur f entraîne que g ou h est un élément de $K_A^\times$, disons g. Mais $g \in A[X]$, et ainsi $g \in A \setminus \{0\}$. En particulier, g est un diviseur commun à tous les coefficients de f. Le polynôme f étant primitif, on obtient $g \in A^\times$. Ainsi, f est irréductible dans l'anneau $A[X]$, et cela achève la démonstration. □

La question restant ouverte pour le moment est donc la suivante : un polynôme unitaire irréductible dans $A[X]$ est-il irréductible dans $K_A[X]$?

Hélas, trois fois hélas, la réponse est négative en général.

2.18. Exemple. Si $A = \mathbb{Z}[\sqrt{5}]$, alors $K_A = \mathbb{Q}[\sqrt{5}] = \{a + b\sqrt{5} \mid a, b \in \mathbb{Q}\}$. Le polynôme $f = X^2 - X - 1 \in A[X]$ n'est pas irréductible dans $K_A[X]$, car on a

$$f = \Big(X - \frac{1+\sqrt{5}}{2}\Big)\Big(X - \frac{1-\sqrt{5}}{2}\Big) \in K_A[X].$$

En revanche, il est malgré tout irréductible dans $A[X]$. En effet, f est non nul, non constant, donc en particulier non inversible dans $A[X]$. Si de plus on a $f = gh$, $g, h \in A[X]$, en comparant les coefficients dominants, on constate que les coefficients dominants de g et h sont inversibles, et inverses l'un de l'autre. On peut donc supposer que g et h sont unitaires. Mais alors, g est un diviseur unitaire de f dans $K_A[X]$. La décomposition précédente montre alors que

$$g = 1,\ f,\ X - \frac{1+\sqrt{5}}{2} \quad \text{ou} \quad X - \frac{1-\sqrt{5}}{2}.$$

Les deux derniers polynômes n'étant pas à coefficients dans A, on obtient donc $g = 1$ ou f. Dans le second cas, $h = 1$. Ainsi, f est bien irréductible dans $A[X]$.

Nous allons montrer dans la suite que si A est factoriel, un polynôme non constant irréductible dans $A[X]$ est encore irréductible dans $K_A[X]$. En particulier, les critères donnés dans ce paragraphe montrent également l'irréductibilité de f dans $K_A[X]$ lorsque A est factoriel.

En fait, nous verrons dans le chapitre XXVI que la bonne notion pour ces considérations est la notion d'anneau intégralement clos.

3. Factorialité de $A[X]$

On s'intéresse ici aux conditions sous lesquelles $A[X]$ est un anneau factoriel.

Nous allons tout d'abord nous intéresser aux éléments irréductibles de $A[X]$ lorsque A est factoriel.

Si A est intègre, on sait déjà qu'un élément $\pi \in A$ est irréductible dans l'anneau A si, et seulement si, il l'est dans $A[X]$ (cf. lemme 1.1). Ce simple résultat a déjà une conséquence intéressante.

3.1. Corollaire. *Si $A[X]$ est factoriel, A est factoriel.*

Démonstration. Supposons $A[X]$ factoriel. Alors, $A[X]$ est intègre, et donc A est intègre par le lemme XV-6.9. Soit $a \in A \setminus \{0\}$. Puisque $A[X]$ est factoriel, on peut écrire

$$a = uf_1 \cdots f_r,$$

où $u \in A[X]^\times$ et f_i est irréductible pour tout $i \in [\![1, r]\!]$. Comme A est intègre, on a

$$\deg(a) = 0 = \deg(u) + \deg(f_1) + \cdots + \deg(f_r),$$

toujours par le lemme XV-6.9. On en déduit que $u, f_1, \ldots, f_r \in A$. Par le lemme 1.1, $f_1, \ldots, f_r$ sont irréductibles dans A. Enfin, on a $u \in A^\times$, car $A[X]^\times = A^\times$ par le lemme XV-6.31. Ainsi, a possède une décomposition en produit d'éléments irréductibles de A et d'une unité. L'unicité de la décomposition provient du fait qu'une décomposition dans A est aussi une décomposition dans $A[X]$ (puisque qu'un élément irréductible de A reste irréductible dans $A[X]$), et de l'unicité de la décomposition dans $A[X]$. □

Avant de continuer, nous allons définir le contenu d'un polynôme de $K_A[X]$. On commence par étendre la notion d'élément normalisé au corps des fractions. Au vu de la remarque XVIII-5.13 (2), la définition suivante a un sens.

3.2. Définition. Soit A un anneau factoriel, soit K_A son corps des fractions et soit $\mathscr{P}$ un s.c.r.i. de A. On dit que $x \in K_A$ est *normalisé* (par rapport à $\mathscr{P}$) s'il est nul, ou s'il est non nul et s'écrit sous la forme

$$x = \prod_{\pi \in \mathscr{P}} \pi^{n_\pi},$$

avec $n_\pi \in \mathbb{Z}$, où les entiers n_π sont nuls sauf un nombre fini.

Tout élément de $K_A^\times$ s'écrit donc de manière unique comme produit d'un élément inversible et d'un élément normalisé.

3.3. Remarques

(1) Le produit de deux éléments normalisés est normalisé. De plus, si $a \in A$ et $b \in A \setminus \{0\}$ sont normalisés, alors $\dfrac{a}{b}$ est normalisé.

(2) Tout élément normalisé $x \in K_A^\times$ peut s'écrire $x = \dfrac{a}{b}$, où a et b sont normalisés.

En effet, si $x = \prod_{\pi \in \mathscr{P}} \pi^{n_\pi}$, il suffit de poser

$$a = \prod_{\pi \in \mathscr{P}, n_\pi > 0} \pi^{n_\pi} \quad \text{et} \quad b = \prod_{\pi \in \mathscr{P}, n_\pi < 0} \pi^{-n_\pi}.$$

On continue par un lemme.

3.4. Lemme. *Soit A un anneau factoriel, et soit $\mathscr{P}$ un s.c.r.i. de A. Pour tout polynôme non nul $f \in K_A[X]$, il existe un unique élément $c \in K_A^\times$ normalisé et un unique polynôme $f_1 \in A[X]$ primitif tels que $f = cf_1$.*

Démonstration. Soit b un dénominateur commun des coefficients de f, que l'on peut toujours supposer normalisé, quitte à multiplier par un élément inversible. Alors, $bf \in A[X]$. Soit alors a le pgcd normalisé des coefficients de bf. On peut donc écrire $bf = af_1$, où $f_1 \in A[X]$ est un polynôme primitif. On pose alors $c = \dfrac{a}{b}$ (qui est donc normalisé), et l'on a $f = cf_1$.

Supposons que $f = cf_1 = dg_1$, où $c, d \in K_A^\times$ sont normalisés et f_1, g_1 sont primitifs. On écrit

$$c = \frac{a}{b}, \quad d = \frac{a'}{b'},$$

avec a, a', b, b' normalisés. Ainsi, ab' et $a'b$ sont aussi normalisés. En comparant les pgcd normalisés des coefficients de $ab'f_1$ et $a'bg_1$, qui sont tous deux égaux à $bb'f_1$, on en déduit que $ab' = a'b$, d'où $c = d$ et $f_1 = g_1$. Cela achève la démonstration. □

3.5. Définition. Soit A un anneau factoriel, et soit $\mathscr{P}$ un s.c.r.i. de A. Si $f \in K_A[X]$ est un polynôme non nul, le *contenu* (normalisé) de f, est l'unique élément normalisé $c \in K_A^\times$ tel que $f = cf_1$, où $f_1 \in A[X]$ est primitif. On le note $c(f)$.

3.6. Lemme. *Soit A un anneau factoriel, et soit $\mathscr{P}$ un s.c.r.i. de A. Enfin, soit $f \in K_A[X]$ non nul. Alors :*

(1) *pour tout $\alpha \in K_A^\times$ normalisé, on a $c(\alpha f) = \alpha c(f)$;*

(2) *on a $f \in A[X]$ si, et seulement si, $c(f) \in A \setminus \{0\}$. Dans ce cas, $c(f)$ est le pgcd normalisé des coefficients de f ;*

(3) *on a $f \in A[X]$ et f est primitif si, et seulement si, $c(f) = 1$.*

Démonstration

(1) Par définition, $f = c(f)f_1$, avec $f_1 \in A[X]$ primitif. Si $\alpha \in K_A^\times$ est normalisé, alors $\alpha f = \big(\alpha c(f)\big)f_1$. Or, $\alpha c(f)$ est normalisé, d'où le résultat.

(2) Si $f \in A[X]$, et si c est le pgcd normalisé des coefficients de f, alors il est clair que $f = cf_1$, où $f_1 \in A[X]$ est primitif. Ainsi, $c(f) = c \in A \setminus \{0\}$. Inversement, si $c(f) \in A \setminus \{0\}$, alors $f = c(f)f_1 \in A[X]$, puisque $f_1 \in A[X]$.

(3) Si $f \in A[X]$ et est primitif, alors le pgcd normalisé de ses coefficients est 1, et l'on a ainsi $c(f) = 1$, par le point précédent.

Inversement, si $c(f) = 1$, alors

$$f = c(f)f_1 = f_1.$$

Ainsi, $f \in A[X]$ et f est primitif. □

On continue par un lemme dû à Gauss.

3.7. Lemme (Gauss). *Soit A un anneau factoriel. Alors, pour tous polynômes $f, g \in K_A[X]$ non nuls, on a $c(fg) = c(f)c(g)$.*

Démonstration. Écrivons $f = c(f)f_1$ et $g = c(g)g_1$, où $f_1, g_1 \in A[X]$ sont primitifs. Supposons avoir montré que le polynôme f_1g_1 est primitif. Alors, puisque $fg = c(f)c(g)f_1g_1$ et $c(f)c(g) \in K_A^\times$ est normalisé, on a

$$c(fg) = c(f)c(g).$$

Il reste donc à montrer que si deux polynômes $f, g \in A[X]$ sont primitifs, alors fg est aussi primitif.

Supposons au contraire que π soit un élément irréductible qui divise tous les coefficients de fg, et considérons le morphisme de réduction modulo π

$$\begin{aligned} A[X] &\longrightarrow A/(\pi)[X] \\ h &\longmapsto \overline{h}. \end{aligned}$$

Par hypothèse, les coefficients de f et g ne sont pas tous divisibles par π, et donc $\overline{f} \neq \overline{0}$, ainsi que $\overline{g} \neq \overline{0}$. En revanche, tous les coefficients de fg sont divisibles par π, et donc $\overline{fg} = 0$. On a alors

$$\overline{f}\,\overline{g} = \overline{fg} = \overline{0}.$$

Mais π est irréductible, donc (π) est un idéal premier par le lemme d'Euclide. Ainsi, $A/(\pi)$ est intègre et donc $\big(A/(\pi)\big)[X]$ est aussi intègre, d'où une contradiction. Par conséquent, fg est primitif, et l'on a fini. □

3.8. Corollaire. *Soit A un anneau factoriel, et soient $f, g, h \in A[X]$ non nuls, tels que $f = gh$. Alors, f est primitif si, et seulement si, g et h sont primitifs.*

Démonstration. Puisque $f, g, h, \in A[X]$, on a $c(f), c(g), c(h) \in A \setminus \{0\}$ par le lemme 3.6 (2). Comme $c(f) = c(g)c(h)$, on a

$$c(f) = 1 \iff c(g)c(h) = 1 \iff c(g), c(h) \in A^\times.$$

Comme $c(g)$ et $c(h)$ sont normalisés, on en déduit

$$c(f) = 1 \iff c(g) = c(h) = 1.$$

Puisqu'un polynôme non nul de $A[X]$ est primitif si, et seulement si, son contenu est 1 par le lemme 3.6 (3), on a fini. □

3.9. Remarque. Le lemme précédent et son corollaire sont faux si A est un anneau quelconque (même intègre). Un contre-exemple est proposé au lecteur dans l'exercice 13.

On peut maintenant décrire les éléments irréductibles de $A[X]$ non constants.

3.10. Lemme. *Soit A un anneau factoriel, et soit $f \in A[X]$ un polynôme non constant. Alors, f est irréductible si, et seulement si, il est primitif et irréductible dans $K_A[X]$.*

Démonstration. Supposons que f soit primitif, et irréductible dans $K_A[X]$. Alors, il est irréductible dans $A[X]$ d'après le lemme 2.17, puisqu'un anneau factoriel est intègre.

Inversement, supposons que $f \in A[X]$ soit irréductible. On a déjà remarqué que f est alors primitif. Il reste à montrer que f est irréductible dans l'anneau $K_A[X]$. Comme f est non constant, il est non nul et non inversible dans $K_A[X]$. Supposons que $f = gh$, $g, h \in K_A[X]$. Alors, d'après le lemme 3.7, on a $c(g)c(h) = 1$. Si $c = c(g)$, on a donc

$$g = cg_1, \quad h = c^{-1}h_1,$$

où $g_1, h_1 \in A[X]$ sont primitifs et $c \in K_A^\times$ est normalisé. Ainsi, on a

$$f = g_1 h_1.$$

Comme f est irréductible dans $A[X]$, on en déduit que $g_1 \in A[X]^\times = A^\times$ par exemple. En particulier, $g = cg_1 \in K_A^\times$. Ainsi, f est irréductible dans l'anneau $K_A[X]$. □

3.11. Remarque. Le lemme précédent montre que, si A est factoriel, un polynôme $f \in A[X]$ non constant irréductible dans $A[X]$ est aussi irréductible dans $K_A[X]$.

En particulier, les critères de réduction modulo un idéal ou d'Eisenstein donnent aussi l'irréductibilité dans $K_A[X]$, dès que A est factoriel. Les considérations faites à la fin du paragraphe précédent permettent ainsi d'utiliser ces critères pour démontrer l'irréductibilité d'un polynôme à coefficients dans K_A.

3.12. Exemple. Pour tout $n \geqslant 1$, le polynôme $X^n - 2$ est non constant, et irréductible dans $\mathbb{Z}[X]$ d'après l'exemple 2.16. Il est donc aussi irréductible dans $\mathbb{Q}[X]$.

On peut maintenant énoncer et démontrer le théorème suivant.

3.13. Théorème (Gauss). *Soit A un anneau. Alors, $A[X]$ est un anneau factoriel si, et seulement si, A est factoriel.*

Dans ce cas, les éléments irréductibles de $A[X]$ sont :

(1) *les éléments irréductibles de A ;*

(2) *les polynômes non constants qui sont primitifs et irréductibles dans l'anneau $K_A[X]$.*

Démonstration. Supposons que A soit un anneau factoriel. La dernière partie a alors été démontrée dans les lemmes 1.1 et 3.10. Montrons maintenant que $A[X]$ est factoriel. Notons que $A[X]$ est intègre, car A l'est. Montrons l'existence d'une décomposition d'un élément non nul en produit d'une unité et d'éléments irréductibles. Soit $f \in A[X]$ non nul. On se fixe un s.c.r.i. de A, et l'on écrit $f = c(f)f_1$, où $f_1 \in A[X]$ est primitif.

Comme $f \in A[X]$, $c(f) \in A \setminus \{0\}$, et se décompose donc en produit d'un élément inversible et d'éléments irréductibles de A (qui sont des éléments irréductibles de $A[X]$). Il reste à décomposer f_1. On écrit

$$f_1 = \beta g_1 \cdots g_r,$$

où $\beta \in K_A^\times$ et g_i est irréductible dans $K_A[X]$ (c'est possible, car $K_A[X]$ est factoriel par la remarque XV-6.13). On écrit alors $g_i = c(g_i)h_i$, où le polynôme $h_i \in A[X]$ est primitif. Remarquons que h_i diffère de g_i par un élément de $K_A^\times$, donc h_i est aussi irréductible dans $K_A[X]$. Ainsi, h_i est irréductible dans $A[X]$. On a alors

$$f_1 = \alpha h_1 \cdots h_r, \quad \alpha \in K_A^\times,$$

et par le lemme de Gauss, en calculant les contenus, on obtient $\alpha = 1$. Ainsi, on a bien décomposé f_1, et donc f, en produit d'une unité et d'éléments irréductibles de $A[X]$.

Il reste à démontrer l'unicité d'une telle décomposition. Supposons que l'on ait une égalité

$$u\pi_1 \cdots \pi_r f_1 \cdots f_m = v\pi'_1 \cdots \pi'_s g_1 \cdots g_n,$$

où $u, v \in A^\times$, $\pi_i, \pi'_j \in A$ sont irréductibles dans A (pas nécessairement dans le s.c.r.i. fixé) et $f_k, g_\ell \in A[X]$ sont primitifs et irréductibles dans $K_A[X]$.

Puisque $K_A[X]$ est factoriel et que $u\pi_1 \cdots \pi_r$ et $v\pi'_1 \cdots \pi'_s$ sont inversibles dans l'anneau $K_A[X]$, on obtient $m = n$ et $g_k = u_k f_{\sigma(k)}$ pour tout $k \in [\![1, m]\!]$ et un certaine permutation $\sigma \in \mathfrak{S}_m$, avec $u_k \in K_A^\times$. En passant aux contenus, on voit que $u_k = 1$.

Par intégrité de $A[X]$, on obtient

$$u\pi_1 \cdots \pi_r = v\pi'_1 \cdots \pi'_s.$$

Puisque A est factoriel, $r = s$ et $\pi'_i = v_i \pi_{\tau(i)}$ pour tout entier $i \in [\![1, r]\!]$, avec $\tau \in \mathfrak{S}_r$, et $v_i \in A^\times$. Ainsi, $A[X]$ est factoriel. La réciproque étant donnée par le corollaire 3.1, cela achève la démonstration. □

3.14. Corollaire. *Si A est factoriel, alors $A[X_1, \ldots, X_n]$ est factoriel.*

3.15. Exemples

(1) Puisque $\mathbb{Z}$ est principal, donc factoriel, l'anneau $\mathbb{Z}[X]$ est factoriel. Nous verrons plus loin qu'il est noethérien.

(2) Soit $A = \mathbb{C}[X_1, X_2, \ldots] = \bigcup_{n \geqslant 1} \mathbb{C}[X_1, \ldots, X_n]$, l'anneau des polynômes en une infinité d'indéterminées. Alors, A est un anneau factoriel, non noethérien.

En effet, soit $\mathfrak{a}$ l'idéal engendré par les X_i, $i \geqslant 1$. Nous affirmons que $\mathfrak{a}$ n'est pas de type fini.

Supposons le contraire, et soient $P_1, \ldots, P_n \in A$ des générateurs de $\mathfrak{a}$. Chaque P_j faisant intervenir un nombre fini d'indéterminées, il existe un entier $m \geqslant 1$ tel que $P_j \in \mathbb{C}[X_1, \cdots, X_m]$ pour tout j. Par définition de $\mathfrak{a}$, on a $X_{m+1} \in \mathfrak{a}$, donc par hypothèse, nous avons une relation du type

$$X_{m+1} = Q_1 P_1 + \cdots + Q_n P_n, \quad Q_i \in A.$$

Soit $s \geqslant m + 1$ un entier tel que Q_j, $P_j \in \mathbb{C}[X_1, \ldots, X_s]$ pour tout j. En substituant 1 à X_{m+1}, on obtient

$$1 = R_1 P_1 + \cdots + R_n P_n, \quad R_i \in \mathbb{C}[X_1, \ldots, X_s],$$

puisque $s \geqslant m + 1$ et puisque les polynômes P_i ne font intervenir que les indéterminées $X_1, \ldots, X_m$. On a donc $1 \in \mathfrak{a}$. Or, cela est impossible, car les éléments de $\mathfrak{a}$ n'ont pas de terme constant, d'où une contradiction.

Vérifions que l'anneau intègre A est factoriel. Remarquons tout d'abord que pour tout $n \geqslant 1$, un élément irréductible $f \in \mathbb{C}[X_1, \ldots, X_n]$ est irréductible dans A. En effet, si f est irréductible dans $\mathbb{C}[X_1, \ldots, X_n]$, alors f est non nul et non inversible dans $\mathbb{C}[X_1, \ldots, X_n]$.

Comme on a $A^\times = \mathbb{C}^\times = \mathbb{C}[X_1, \ldots, X_n]^\times$, f est aussi non inversible dans A. Si maintenant $f = gh$, $g, h \in A$, il existe un entier $m \geqslant n$ tel que $f, g, h \in \mathbb{C}[X_1, \ldots, X_m]$. D'après le lemme 1.1, f est aussi irréductible dans l'anneau $\mathbb{C}[X_1, \ldots, X_m]$. Ainsi, g ou h est inversible dans $\mathbb{C}[X_1, \ldots, X_m]$, donc dans A. Par conséquent, f est irréductible dans A.

Soit $f \in A \setminus \{0\}$. Alors, il existe $n \geqslant 1$ tel que $f \in \mathbb{C}[X_1, \ldots, X_n]$. Comme $\mathbb{C}[X_1, \ldots, X_n]$ est factoriel par le corollaire précédent, f se décompose en produit d'éléments irréductibles de $\mathbb{C}[X_1, \ldots, X_n]$ et d'un élément inversible. Les considérations précédentes montrent que cette décomposition fournit aussi une décomposition en irréductibles dans A. De plus, deux décompositions de f peuvent être vues dans un même anneau $\mathbb{C}[X_1, \ldots, X_n]$ pour un entier $n \geqslant 1$ suffisamment grand. L'unicité découle alors encore une fois de la factorialité de $\mathbb{C}[X_1, \ldots, X_n]$.

4. Irréductibilité d'un polynôme bicarré

Nous allons voir maintenant que, bien que les critères d'irréductibilité établis dans ce chapitre soient bien pratiques, ils ont leurs limites. Pour tous entiers a, $b \in \mathbb{Z}$, on pose

$$f_{a,b} = X^4 - 2(a+b)X^2 + (a-b)^2 \in \mathbb{Z}[X].$$

Le but de ce paragraphe est d'étudier l'irréductibilité de ce polynôme dans l'anneau $\mathbb{Q}[X]$.

On commence par éliminer les cas faciles. On remarque tout d'abord que l'on a

$$f_{a,b} = \big(X - \sqrt{a} - \sqrt{b}\big)\big(X - \sqrt{a} + \sqrt{b}\big)\big(X + \sqrt{a} - \sqrt{b}\big)\big(X + \sqrt{a} + \sqrt{b}\big) \in \mathbb{C}[X].$$

Si $a = r^2$, $r \in \mathbb{Q}$, on a donc

$$f_{a,b} = \big((X-r)^2 - b\big)\big((X+r)^2 - b\big),$$

et $f_{a,b}$ n'est pas irréductible. Il en est de même si b est un carré dans $\mathbb{Q}$.

Enfin, si $b = ar^2$, $r \in \mathbb{Q}$, on a

$$f_{a,b} = \big(X^2 - a(1+r)^2\big)\big(X^2 - a(1-r)^2\big),$$

et $f_{a,b}$ n'est pas irréductible.

On suppose dorénavant que a et b sont non nuls, et que a, b, $\frac{b}{a} \notin \mathbb{Q}^{\times 2}$.

Soient $u \in \mathbb{Z}$, $v \in \mathbb{Z} \setminus \{0\}$. Par le lemme 1.6, f est irréductible dans l'anneau $\mathbb{Q}[X]$ si, et seulement si, $f_{a,b}\big(X + \frac{u}{v}\big)$ est irréductible, ce qui équivaut encore au fait que $f_{a,b}(vX + u)$ est irréductible, en multipliant par une puissance de v convenable.

Un simple calcul montre alors le résultat suivant.

4.1. Lemme. *Le polynôme $f_{a,b}(vX+u)$ est égal à*

$$v^4X^4 + 4uv^3X^3 + 2v^2\big(3u^2 - (a+b)\big)X^2 + 4uv\big(u^2 - (a+b)\big)X + u^4 - 2(a+b)u^2 + (a-b)^2.$$

Supposons maintenant que l'on veuille trouver un changement de variables du type $X \longleftrightarrow vX+u$ de manière à pouvoir appliquer le critère de réduction modulo un nombre premier.

En particulier, on doit avoir $p \nmid v$ à cause de la condition sur le coefficient dominant. Ainsi, $\overline{v}$ est inversible dans $\mathbb{F}_p$.

4.2. Proposition. *Soit p un nombre premier, et soient $u, v \in \mathbb{Z}$, $p \nmid v$. Alors, $f_{a,b}(vX+u)$ est réductible modulo p.*

En particulier, $\overline{f}$ est réductible modulo tout nombre premier.

Démonstration. Soit $g = f(vX+u)$. Alors, $\overline{g} = \overline{f}(\overline{v}X + \overline{u})$, et $\overline{g}$ est irréductible si, et seulement si, $\overline{f}(X + \overline{u}\,\overline{v}^{-1})$ est irréductible, ou encore si, et seulement si, $\overline{f}$ est irréductible par le lemme 1.6.

Il suffit donc de vérifier que $\overline{f}$ est réductible modulo p. On remarque que

$$f_{a,b} = \big(X^2 - (a+b)\big)^2 - 4ab.$$

Si $p \mid 4ab$, on a donc

$$\overline{f} = (X^2 - \overline{a+b})^2,$$

et $\overline{f}$ n'est pas irréductible.

Supposons que $p \nmid 4ab$. En particulier, p est impair et $\overline{a}, \overline{b} \in \mathbb{F}_p^\times$.

▷ Si $\overline{a} \in \mathbb{F}_p^{\times 2}$, on vérifie que l'on a

$$\overline{f} = \big((X-\overline{w})^2 - \overline{b}\big)\big((X+\overline{w})^2 - \overline{b}\big),$$

où $\overline{a} = \overline{w}^2$.

▷ Si $\overline{b} \in \mathbb{F}_p^{\times 2}$, on vérifie que l'on a

$$\overline{f} = \big((X-\overline{w})^2 - \overline{a}\big)\big((X+\overline{w})^2 - \overline{a}\big),$$

où $\overline{b} = \overline{w}^2$.

▷ Si $\overline{a}, \overline{b} \notin \mathbb{F}_p^{\times 2}$, comme $\mathbb{F}_p^\times/\mathbb{F}_p^{\times 2}$ est d'ordre 2 par le lemme XVII-4.1, $\overline{a}$ et $\overline{b}$ représentent la même classe de carrés, et donc $\overline{a}\,\overline{b} \in \mathbb{F}_p^{\times 2}$.

On vérifie alors que l'on a

$$\overline{f} = (X^2 - \overline{a+b} - \overline{2w})(X^2 - \overline{a+b} + \overline{2w}),$$

où $\overline{ab} = \overline{w}^2$. □

Supposons maintenant que l'on veuille trouver un changement de variables du type $X \longleftrightarrow vX + u$ de manière à pouvoir appliquer le critère d'Eisenstein. Encore une fois, on doit avoir $p \nmid v$.

4.3. Proposition. *Soit p un nombre premier, et soient $u, v \in \mathbb{Z}$, $p \nmid v$. Alors, $f_{a,b}(vX + u)$ n'est pas un polynôme d'Eisenstein en p.*

Démonstration. Par le lemme 4.1, $f_{a,b}(vX + u)$ est égal à

$$v^4X^4 + 4uv^3X^3 + 2v^2\big(3u^2 - (a+b)\big)X^2 + 4uv\big(u^2 - (a+b)\big)X + u^4 - 2(a+b)u^2 + (a-b)^2.$$

Supposons que le polynôme $f_{a,b}(vX + u)$ soit un polynôme d'Eisenstein en p. Alors, $p \mid 4uv^3$, et donc $p = 2$ ou $p \mid u$ (car $p \nmid v$).

Supposons que $p \neq 2$. Alors, $p \mid u$. Mais alors, la condition

$$p \mid u^4 - 2(a+b)u^2 + (a-b)^2$$

implique que $p \mid (a-b)^2$, d'où $p \mid a - b$. Mais alors, comme $p \mid u$, on en déduit que $p^2 \mid u^4 - 2(a+b)u^2 + (a-b)^2$, d'où une contradiction.

Ainsi, $p = 2$. On a

$$u^4 - 2(a+b)u^2 + (a-b)^2 = \big(u^2 - (a+b)\big)^2 - 4ab.$$

Si $u^2 - (a+b)$ est impair, $u^4 - 2(a+b)u^2 + (a-b)^2$ est impair, d'où une contradiction. Si $u^2 - (a+b)$ est pair, $u^4 - 2(a+b)u^2 + (a-b)^2$ est divisible par 4, d'où encore une contradiction. Ainsi, $f_{a,b}(vX+u)$ n'est pas un polynôme d'Eisenstein en p, et l'on a fini. □

Mais alors, comment montrer l'irréductibilité de $f_{a,b}$?

4.4. Proposition. *Soient $a, b \in \mathbb{Z}$. Alors, le polynôme*

$$f_{a,b} = X^4 - 2(a+b)X^2 + (a-b)^2$$

est irréductible dans l'anneau $\mathbb{Q}[X]$ si, et seulement si, a, b sont non nuls et a, b, $\dfrac{b}{a} \notin \mathbb{Q}^{\times 2}$.

Démonstration. On a déjà remarqué que les conditions évoquées sont nécessaires. Montrons qu'elles sont suffisantes. On a

$$f_{a,b} = (X - \sqrt{a} - \sqrt{b})(X - \sqrt{a} + \sqrt{b})(X + \sqrt{a} - \sqrt{b})(X + \sqrt{a} + \sqrt{b}) \in \mathbb{C}[X].$$

Si $g \in \mathbb{Q}[X]$ est un facteur unitaire non trivial de $f_{a,b}$, on peut toujours supposer qu'il est de degré 1 ou 2. Ainsi, il est produit de 1 ou 2 des 4 facteurs ci-dessus.

Fait. Pour tous $\alpha, \beta \in \mathbb{Z}$, $(\alpha, \beta) \neq (0,0)$, $x = \alpha\sqrt{a} + \beta\sqrt{b} \notin \mathbb{Q}$.

En effet, supposons que x soit rationnel. Si α ou β est nul, alors cela entraîne que a ou b est un carré dans $\mathbb{Q}$, d'où une contradiction. Si α et β sont tous deux non nuls, on a $x^2 = a+b+2\alpha\beta\sqrt{a}\sqrt{b} \in \mathbb{Q}$, et l'on obtient que $\sqrt{a}\sqrt{b} \in \mathbb{Q}$. Mais alors, $ab \in \mathbb{Q}^{\times 2}$ et donc $\dfrac{b}{a} = \dfrac{b^2}{ab} \in \mathbb{Q}^{\times 2}$, d'où une contradiction.

Cela démontre que g est nécessairement le produit de deux facteurs de degré 1. En examinant les six cas possibles, on voit que le terme en X est de la forme $m\sqrt{a}$ ou $m\sqrt{b}$, $m \in \mathbb{Z}$. Lorsque $m \neq 0$, aucun de ces nombres complexes n'est rationnel, d'où une contradiction. Il reste le cas $m = 0$, qui correspond aux deux possibilités

$$g = (X - \sqrt{a} - \sqrt{b})(X + \sqrt{a} + \sqrt{b}) \quad \text{ou} \quad (X - \sqrt{a} + \sqrt{b})(X + \sqrt{a} - \sqrt{b}).$$

Mais alors, le terme constant est $-(\sqrt{a} \pm \sqrt{b})^2 = -(a+b \pm 2\sqrt{a}\sqrt{b})$. Encore une fois, ce nombre complexe n'est pas rationnel. Ainsi, le polynôme $f_{a,b}$ est bien irréductible dans $\mathbb{Q}[X]$. □

L'exemple suivant montre l'intérêt de garder en tête les critères d'irréductibilité sous leur forme la plus générale.

4.5. Exemple. Examinons le cas $a = 2$ et $b = 3$. On a donc

$$f_{2,3} = X^4 - 10X^2 + 1.$$

Alors, $f_{2,3}(vX + u)$ ne vérifie pas les hypothèses du critère de réduction modulo p, et cela pour tout nombre premier p (soit $p \mid v$ et c'est fini, soit $p \nmid v$ et l'on utilise la proposition 4.2).

En revanche, nous allons maintenant retrouver le fait que $f_{2,3}$ est irréductible dans $\mathbb{Q}[X]$ en réduisant modulo 9. Puisque $\mathbb{Z}$ est factoriel, il suffit de montrer que $f_{2,3}$ est irréductible dans $\mathbb{Z}[X]$ d'après la remarque 3.11. Pour démontrer ce dernier point, il suffit de démontrer que $\overline{f}_{2,3} \in \mathbb{Z}/9\mathbb{Z}[X]$ est irréductible, d'après la remarque 2.5 (1).

Le polynôme $\overline{f}_{2,3}$ est non nul. Si $\overline{f}_{2,3}$ était inversible, alors la réduction $\tilde{f}_{2,3}$ de $f_{2,3}$ modulo 3 serait inversible. Mais $\tilde{f}_{2,3} = (X^2 + \tilde{1})^2 \in \mathbb{F}_3[X]$, donc $\tilde{f}$ est non inversible dans $\mathbb{F}_3[X]$ (car il est non constant), et $\overline{f}$ est donc non inversible.

Supposons que $\overline{f}_{2,3} = \overline{g}\overline{h} \in \mathbb{Z}/9\mathbb{Z}[X]$. Cela induit une décomposition

$$\tilde{f}_{2,3} = \tilde{g}\tilde{h} \in \mathbb{F}_3[X].$$

Mais on a $\tilde{f}_{2,3} = (X^2 + \tilde{1})^2 \in \mathbb{F}_3[X]$. On a donc

$$\tilde{g} = \tilde{a}, \quad \tilde{a}(X^2 + \tilde{1}) \quad \text{ou} \quad \tilde{a}(X^2 + \tilde{1})^2, \qquad \tilde{a} \in \mathbb{F}_3^{\times}.$$

Puisque $\tilde{a} \in \mathbb{F}_3^\times$, on a $3 \nmid a$. On en déduit que $\overline{a} \in (\mathbb{Z}/9\mathbb{Z})^\times$. Quitte à remplacer $\overline{g}$ par $\overline{a}^{-1}\,\overline{g}$, et $\overline{h}$ par $\overline{a}\,\overline{h}$, on peut supposer que

$$\tilde{g} = \tilde{1}, \quad X^2 + \tilde{1} \quad \text{ou} \quad (X^2 + \tilde{1})^2.$$

Remarquons que si $\tilde{g} = (X^2 + \tilde{1})^2$, alors $\tilde{h} = \tilde{1}$, et donc par symétrie du problème, on peut se restreindre aux deux premiers cas.

(i) On suppose que $\tilde{g} = \tilde{1}$. On a donc

$$\overline{g} = \overline{1} + \overline{3}\,\overline{P}.$$

Mais alors, $\overline{g}$ est inversible puisque

$$(\overline{1} + \overline{3}\,\overline{P})(\overline{1} - \overline{3}\,\overline{P}) = \overline{1}.$$

(ii) On suppose que $\tilde{g} = X^2 + \tilde{1}$. Alors, $\tilde{h} = X^2 + \tilde{1}$. On a donc

$$\overline{g} = X^2 + \overline{1} + \overline{3}\,\overline{P}, \quad \overline{h} = X^2 + \overline{1} + \overline{3}\,\overline{Q}.$$

Mais alors, on a

$$X^4 - X^2 + \overline{1} = (X^2 + \overline{1})^2 + \overline{3}(X^2 + \overline{1})\overline{P + Q} = X^4 + \overline{2}X^2 + \overline{1} + \overline{3}(X^2 + \overline{1})\overline{P + Q},$$

et l'on obtient $\overline{3}\,\overline{X^2 + (X^2 + 1)(P + Q)} = \overline{0}$. On a donc

$$X^2 + (X^2 + 1)(P + Q) = 3S.$$

Ainsi, $X^2 + \tilde{1} \mid -X^2$ dans $\mathbb{F}_3[X]$, d'où une contradiction.

Finalement, si $\overline{f}_{2,3} = \overline{g}\,\overline{h}$, on a montré que $\overline{g}$ ou $\overline{h}$ est inversible. Ainsi, $f_{2,3}$ est irréductible dans $\mathbb{Z}[X]$.

On retrouve ainsi le fait que $f_{2,3}$ est irréductible en utilisant le critère général de réduction, mais pas son corollaire.

Hélas, trois fois hélas, les choses ne se passent pas toujours aussi bien. Nous allons en effet exhiber une famille de polynômes $f_{a,b}$ qui sont réductibles modulo tout entier $m \geqslant 2$. Mieux, nous allons donner une condition nécessaire et suffisante sur a et b pour que $f_{a,b}$ soit réductible modulo tout entier $m \geqslant 2$. Les arguments présentés ici sont une réécriture largement étendue de ceux présentés dans [25], dont est tiré (à la reformulation près) notre théorème 4.14.

On commence par un lemme.

4.6. Lemme. *Un polynôme $f \in \mathbb{Z}[X]$ est réductible modulo tout $m \geqslant 2$ si, et seulement si, il est réductible modulo p^r pour tout p premier et tout entier $r \geqslant 1$.*

Démonstration. Supposons que f soit réductible modulo p^r pour tout p premier et tout $r \geqslant 1$. Soit $m \geqslant 2$. Écrivons $m = p_1^{r_1} \cdots p_r^{r_s}, r_i \geqslant 1$ et où les p_i sont des nombres premiers distincts deux à deux.

Remarquons maintenant que si n_1 et n_2 sont deux entiers premiers entre eux, alors les idéaux $n_1\mathbb{Z}[X]$ et $n_2 \in \mathbb{Z}[X]$ sont comaximaux. En effet, par le théorème de Bézout, il existe u, $v \in \mathbb{Z}$ tels que $un_1 + vn_2 = 1$. On conclut en regardant cette égalité dans $\mathbb{Z}[X]$. De plus, on a

$$n_1\mathbb{Z}[X] \cap n_2\mathbb{Z}[X] = n_1n_2\mathbb{Z}[X].$$

Il s'ensuit que les idéaux $p_i^{r_i}\mathbb{Z}[X]$, $i \in [\![1, s]\!]$ sont deux à deux comaximaux, et que leur intersection est $m\mathbb{Z}[X]$. En utilisant le théorème chinois et l'isomorphisme canonique $\mathbb{Z}[X]/n\mathbb{Z}[X] \simeq \mathbb{Z}/n\mathbb{Z}[X]$ induit par la réduction modulo n, on a un isomorphisme d'anneaux

$$\begin{array}{rccc} \varphi : & \mathbb{Z}/m\mathbb{Z}[X] & \longrightarrow & \mathbb{Z}/p_1^{r_1}\mathbb{Z}[X] \times \cdots \times \mathbb{Z}/p_r^{r_s}\mathbb{Z}[X] \\ & \overline{f} & \longmapsto & (\overline{f}_1, \ldots, \overline{f}_r), \end{array}$$

où $\overline{f}$ est la réduction de f modulo m et $\overline{f}_i$ est la réduction de f modulo $p_i^{r_i}$. Par le lemme XVIII-1.24, il suffit de montrer que $(\overline{f}_1, \ldots, \overline{f}_r)$ est réductible.

Supposons qu'il y ait au moins deux indices i et j pour lesquels $\overline{f}_i$ et $\overline{f}_j$ soient non inversibles, disons $i = 1$ et $j = 2$ pour simplifier l'écriture. Alors, on a

$$(\overline{f}_1, \overline{f}_2, \ldots, \overline{f}_r) = (\overline{f}_1, \overline{1}, \ldots, \overline{1}) \cdot (\overline{1}, \overline{f}_2, \ldots, \overline{f}_r).$$

Aucun des éléments de ce produit n'est inversible, puisqu'ayant par hypothèse au moins une coordonnée non inversible. Ainsi, $(\overline{f}_1, \ldots, \overline{f}_r)$ est réductible dans ce cas.

Supposons qu'il y ait au plus un indice i pour lequel $\overline{f}_i$ soit non inversible, disons $i = 1$. Alors, puisque $\overline{f}_1$ est réductible, on a $\overline{f}_1 = \overline{g}\,\overline{h}$, avec $\overline{g}$ et $\overline{h}$ non inversibles (même si $\overline{f}$ est nul, puisque l'on peut prendre alors $\overline{g}$ et $\overline{h}$ nuls. Dans ce cas, on a

$$(\overline{f}_1, \overline{f}_2, \ldots, \overline{f}_r) = (\overline{g}, \overline{f}_2, \ldots, \overline{f}_r) \cdot (\overline{h}, \overline{f}_2, \ldots, \overline{f}_r).$$

Comme précédemment, aucun des facteurs n'est inversible, et $(\overline{f}_1, \ldots, \overline{f}_r)$ est réductible.

Enfin, si $\overline{f}_i$ est inversible pour tout $i \in [\![1, s]\!]$, alors $(\overline{f}_1, \ldots, \overline{f}_r)$ est inversible et par suite, $\overline{f}$ est inversible, donc réductible. Cela achève la démonstration. □

4.7. Remarque. La démonstration du lemme précédent montre en fait le résultat suivant. Soit $m \geqslant 2$ un entier **fixé**, et soit $m = p_1^{r_1} \times \cdots \times p_r^{r_s}$ sa décomposition en facteurs premiers. Si $f \in \mathbb{Z}[X]$ est réductible modulo $p_i^{r_i}$ pour tout $i \in [\![1, s]\!]$, alors f est réductible modulo m.

La réciproque est fausse. En effet, le polynôme $-2X$ est réductible modulo 6, car $-\overline{2}$ et X ne sont pas inversibles dans $\mathbb{Z}/6\mathbb{Z}[X]$, mais $-2X$ est irréductible modulo 3, puisque $\mathbb{Z}/3\mathbb{Z}$ est un corps et $-2X$ est de degré 1 modulo 3.

On peut même trouver un contre-exemple unitaire. Il suffit de considérer le polynôme X, qui est réductible modulo 6, car $\mathbb{Z}/6\mathbb{Z}$ possède un idempotent non trivial (cf. théorème 1.7). En revanche, il est irréductible modulo 2 et modulo 3.

On va maintenant voir que le polynôme $f_{a,b}$ est réductible modulo p^r, pour presque tout p.

4.8. Lemme. *Soit p un nombre premier et soit $r \geqslant 1$. Soient a, $b \in \mathbb{Z}$ deux entiers tels qu'au moins une des conditions suivantes soit vérifiée :*

(1) *l'équation $x^2 \equiv a\ [p^r]$ a une solution ;*

(2) *l'équation $x^2 \equiv b\ [p^r]$ a une solution ;*

(3) *l'équation $x^2 \equiv ab\ [p^r]$ a une solution.*

Alors, $f_{a,b} = X^4 - 2(a+b)X^2 + (a-b)^2$ est réductible modulo p^r.

Démonstration. Si $u \in \mathbb{Z}$ vérifie $u^2 \equiv a[p^r]$, alors on a

$$f \equiv \big(X^2 + 2uX + (a-b)\big)\big(X^2 - 2uX + (a-b)\big) \mod p^r\mathbb{Z}[X].$$

Si $u \in \mathbb{Z}$ vérifie $u^2 \equiv b[p^r]$, alors on a

$$f \equiv \big(X^2 + 2uX - (a-b)\big)\big(X^2 - 2uX - (a-b)\big) \mod p^r\mathbb{Z}[X].$$

Si $u \in \mathbb{Z}$ vérifie $u^2 \equiv ab[p^r]$, alors on a

$$f \equiv (X^2 - (a+b) + 2u)(X^2 - (a+b) - 2u) \mod p^r\mathbb{Z}[X].$$

Cela achève la démonstration, car un polynôme unitaire non constant n'est pas inversible. □

4.9. Corollaire. *Soient a, $b \in \mathbb{Z}$, et soit p un nombre premier impair tel que $p \nmid a$ et $p \nmid b$. Alors, $f_{a,b}$ est réductible modulo p^r pour tout $r \geqslant 1$.*

Démonstration. On a déjà vu que le groupe $\mathbb{F}_p^\times/\mathbb{F}_p^{\times 2}$ a deux éléments. On en déduit qu'au moins un des entiers a, b ou ab est un carré modulo p, donc un carré modulo p^r pour tout $r \geqslant 1$ par le lemme XVII-3.2. On utilise alors le lemme précédent pour conclure. □

On peut maintenant démontrer la proposition suivante.

4.10. Proposition. *Soient p, ℓ deux nombres premiers impairs distincts tels que*

(1) $p \equiv 1\ [8]$;

(2) $\overline{\ell} \in \mathbb{F}_p^{\times 2}$.

Alors, le polynôme $f_{p,\ell} \in \mathbb{Z}[X]$ est irréductible, mais est réductible modulo tout entier $m \geqslant 2$.

Démonstration. Le polynôme $f_{p,\ell}$ est irréductible, car p et ℓ vérifient les conditions de la proposition 4.4. Montrons qu'il est réductible modulo tout entier $m \geqslant 2$. Par le lemme 4.6, il suffit de vérifier que $f_{p,\ell}$ est réductible modulo toute puissance d'un nombre premier. Par le corollaire précédent, il reste à montrer $f_{a,b}$ est réductible modulo 2^r, modulo p^r et modulo ℓ^r pour tout $r \geqslant 1$. Par le lemme XVII-3.2, la condition (1) montre que p est un carré modulo 2^r pour tout $r \geqslant 1$. Par ce même lemme, on voit que la condition (2) montre que ℓ est un carré modulo p^r pour tout $r \geqslant 1$. La loi de réciprocité quadratique (théorème XVII-4.6) et les conditions (1) et (2) montrent que $\overline{p} \in \mathbb{F}_\ell^{\times 2}$. Mais alors, p un carré modulo ℓ^r pour tout entier $r \geqslant 1$. On conclut alors par le lemme 4.8. □

4.11. Exemple. Le polynôme $f_{17,19} = X^4 - 72X^2 + 4 \in \mathbb{Z}[X]$ est irréductible, mais est réductible modulo tout entier $m \geqslant 2$.

En effet, on a $17 \equiv 1\ [8]$ et $19 \equiv 2 \equiv 6^2\ [17]$.

Pour finir, nous allons donner une condition nécessaire et suffisante sur a et b pour que $f_{a,b}$ soit réductible modulo tout entier $m \geqslant 2$.

Auparavant, nous avons besoin de résultats techniques.

4.12. Lemme. *Soit p un nombre premier et soit $r \geqslant 1$. Soit $F \in \mathbb{Z}/p^r\mathbb{Z}[X]$. On suppose que $F \notin (\overline{p})$. Alors, F s'écrit de manière unique*

$$F = UF_0,$$

où $U \in (\mathbb{Z}/p^r\mathbb{Z}[X])^\times$ et $F_0 \in \mathbb{Z}/p^r\mathbb{Z}[X]$ est unitaire.

Démonstration. Soit $f \in \mathbb{Z}[X]$ tel que $F = \overline{f}$. Par hypothèse sur F, la réduction $\tilde{f}$ de f modulo p est non nulle.

Soit $\tilde{v} \in \mathbb{Z}/p\mathbb{Z}$ le coefficient dominant de $\tilde{f}$. Alors, $p \nmid v$ et ainsi v est inversible modulo p^r. Soit $\overline{u} \in \mathbb{Z}/p^r\mathbb{Z}$ l'inverse de $\overline{v} \in \mathbb{Z}/p^r\mathbb{Z}$.

On va montrer qu'il existe des r-uplets $(f_k)_{k\in[\![1,r]\!]}$ et $(g_k)_{k\in[\![1,r]\!]}$ de polynômes de $\mathbb{Z}[X]$ tels que

(1) le polynôme $\overline{f}_k \in \mathbb{Z}/p^k\mathbb{Z}[X]$ est unitaire, pour tout $k \in [\![1, r]\!]$;

(2) on a $f_{k+1} \equiv f_k\ [p^k]$, pour tout $k \in [\![1, r-1]\!]$;

(3) on a $uf \equiv (1 + pg_k)f_k\ [p^k]$, pour tout $k \in [\![1, r]\!]$.

On remarque que u étant un inverse de v modulo p^r, c'est aussi un inverse de v modulo p. On peut donc poser $f_1 = uf$ et $g_1 = 0$. Supposons maintenant les polynômes f_k et g_k construits pour un certain entier $k \in [\![1, r-1]\!]$. Il existe donc $h_k \in \mathbb{Z}[X]$ tel que

$$uf = (1 + pg_k)f_k + p^k h_k.$$

Par hypothèse, $\overline{f}_k \in \mathbb{Z}/p^k\mathbb{Z}[X]$ est unitaire. Par le théorème XV-6.10, il existe donc $Q, R \in \mathbb{Z}/p^k\mathbb{Z}[X]$, avec $\deg(R) < \deg(\overline{f}_k)$ tels que

$$\overline{h}_k = Q\overline{f}_k + R.$$

On choisit q, $r \in \mathbb{Z}[X]$ tels que $Q = \overline{q}, R = \overline{r}$, et $\deg(R) = \deg(r)$.

Il existe alors $h_{k+1} \in \mathbb{Z}[X]$ tel que $h_k = qf_k + r + p^k h_{k+1}$. Posons

$$f_{k+1} = f_k + p^k r \quad \text{et} \quad g_{k+1} = g_k + p^{k-1}q.$$

Par définition, on a donc $f_{k+1} \equiv f_k\ [p^k]$. Remarquons que l'on a

$$\deg(p^k r) = \deg(r) = \deg(R) < \deg(\overline{f}_k) \leqslant \deg(f_k),$$

ce qui implique que f_{k+1} et f_k ont même coefficient dominant. Ainsi, le polynôme $\overline{f}_{k+1} \in \mathbb{Z}/p^{k+1}\mathbb{Z}[X]$ est unitaire. Enfin, on a

$$(1+pg_{k+1})f_{k+1} = (1+pg_k+p^k q)(f_k+p^k r) \equiv (1+pg_k)f_k + p^k q f_k + p^k r\ [p^{k+1}].$$

Cela se récrit

$$(1 + pg_{k+1})f_{k+1} \equiv (1 + pg_k)f_k + p^k(qf_k + r)\ [p^{k+1}],$$

soit encore

$$(1 + pg_{k+1})f_{k+1} \equiv (1 + pg_k)f_k + p^k(h_k - p^k h_{k+1})\ [p^{k+1}].$$

Ainsi, on a

$$(1 + pg_{k+1})f_{k+1} \equiv (1 + pg_k)f_k + p^k h_k \equiv uf [p^{k+1}],$$

d'où le résultat.

Posons maintenant $U = \overline{v}(\overline{1} + \overline{pg}_r)$ et $F_0 = \overline{f}_r$. Alors, F_0 est unitaire, et U est inversible par la proposition XV-6.37, puisque $\overline{p}$ est nilpotent. De plus, on a $F = \overline{f} = UF_0$ par construction, d'où l'existence d'une décomposition vérifiant les conditions de l'énoncé.

Montrons maintenant l'unicité.

Supposons que l'on ait l'égalité $UF_0 = VG_0$, avec $U, V \in (\mathbb{Z}/p^r\mathbb{Z}[X])^\times$ et où $F_0, G_0 \in \mathbb{Z}/p^r\mathbb{Z}[X]$ sont unitaires. Alors, on a $G_0 = V^{-1}UF_0$. Comme F_0 est unitaire, $V^{-1}U$ et $V^{-1}UF_0$ ont même coefficient dominant. Ainsi, $V^{-1}U$ est unitaire. Mais comme $V^{-1}U$ est inversible, il est constant, car sinon son coefficient dominant serait nilpotent (et donc différent de $\overline{1}$)

par la proposition XV-6.37. Ainsi, le polynôme $V^{-1}U$ est constant, unitaire, donc égal à $\overline{1}$, d'où $U = V$, ainsi que $F_0 = G_0$. Cela achève la démonstration. □

4.13. Corollaire. *Soit $F \in \mathbb{Z}/p^r\mathbb{Z}[X]$ un polynôme unitaire non constant. Si F est réductible, alors F est produit de deux polynômes unitaires non constants.*

Démonstration. Notons que F est non nul et non inversible, car il est non constant et $\overline{1}$ est non nilpotent (cf. proposition XV-6.37). Puisque F n'est pas irréductible, on a donc $F = GH$, où G, $H \in \mathbb{Z}/p^r\mathbb{Z}[X]$ sont non inversibles. Par le lemme précédent, on peut écrire

$$G = UG_0, \quad H = VH_0,$$

où U, V sont inversibles et G_0, H_0 sont unitaires. On a alors

$$F = UVG_0H_0.$$

Comme G_0 et H_0 sont unitaires, G_0H_0 l'est aussi. Comme UV est inversible et que F est unitaire, la partie unicité du même lemme montre que $UV = 1$ et $F = G_0H_0$. Comme G est non inversible, G_0 est non inversible. En particulier, G_0 n'est pas constant (sinon il serait égal à $\overline{1}$, puisque G_0 est unitaire). De même, H_0 n'est pas constant. Cela achève la démonstration. □

Nous sommes maintenant prêts à démontrer le résultat suivant.

4.14. Théorème. *Soient a, $b \in \mathbb{Z}$ et soit p un nombre premier. Alors, le polynôme $f_{a,b} \in \mathbb{Z}[X]$ est réductible modulo p^r pour tout $r \geqslant 1$ si, et seulement si, au moins une des conditions suivantes est vérifiée :*

(1) *pour tout $k \geqslant 1$, a est un carré modulo p^k ;*

(2) *pour tout $k \geqslant 1$, b est un carré modulo p^k ;*

(3) *pour tout $k \geqslant 1$, ab est un carré modulo p^k.*

Démonstration. Si $a = 0$, $b = 0$ ou $a = b$, alors $f_{a,b}$ est le produit de deux polynômes unitaires non constants de $\mathbb{Z}[X]$, donc est réductible modulo p^r pour tout $r \geqslant 1$, et une des trois conditions est toujours vérifiée. On peut donc supposer dans la suite que a, b et $a - b$ sont non nuls.

Le lemme 4.8 montre que si une des trois conditions du lemme est vérifiée, alors $f_{a,b}$ est réductible modulo p^r pour tout $r \geqslant 1$.

Montrons avant de continuer que si $f_{a,b}$ est réductible modulo p^r, alors $\overline{f}_{a,b}$ est le produit de deux polynômes unitaires de degré 2.

D'après le corollaire 4.13, il existe alors deux polynômes unitaires non constants G_0, $H_0 \in \mathbb{Z}/p^r\mathbb{Z}[X]$ tels que $\overline{f}_{a,b} = G_0H_0$. Comme G_0 et H_0 sont unitaires, on a

$$\deg(\overline{f}_{a,b}) = 4 = \deg(G_0H_0) = \deg(G_0) + \deg(H_0).$$

Puisque G_0 est non constant, G_0 est de degré 1, 2 ou 3. Si G_0 est un polynôme de degré 2, H_0 aussi et l'on a fini. Si G_0 est de degré 1, alors on peut écrire $G_0 = X - \overline{m}$. On a alors $f_{a,b}(\overline{m}) = \overline{0}$, et donc

$$\overline{f}_{a,b} = \overline{f}_{a,b} - \overline{f}_{a,b}(\overline{m}) = (X^2 - \overline{m}^2)(X^2 + \overline{m}^2 - \overline{2(a+b)}).$$

Si G_0 est de degré 3, H_0 est de degré 1, et l'on conclut comme précédemment.

On suppose maintenant que $f_{a,b}$ est réductible modulo p^r pour tout $r \geqslant 1$. Écrivons $a - b = p^s t_0$, $s \geqslant 0$, $p \nmid t_0$. Fixons $r \geqslant 3$ tel que $r + s \geqslant v_p(ab) + 4$. Par hypothèse, $f_{a,b}$ est réductible modulo p^{2r+4s}. D'après ce qui précède, le polynôme $\overline{f}_{a,b}$ est produit de deux polynômes unitaires de degré 2. En tenant compte du fait que le coefficient en X^3 est nul, on en déduit l'existence de trois entiers u, v, $w \in \mathbb{Z}$ tels que

$$\overline{f}_{a,b} = (X^2 + \overline{u}X + \overline{v})(X^2 - \overline{u}X + \overline{w}).$$

Par identification, on obtient

$$\begin{aligned} w + v - u^2 &\equiv -2(a+b) \quad [p^{2r+4s}] \qquad (1)\\ u(w-v) &\equiv 0 \quad [p^{2r+4s}] \qquad (2)\\ vw &\equiv (a-b)^2 \quad [p^{2r+4s}]. \qquad (3) \end{aligned}$$

Quitte à ajouter à u, v et w des multiples de p^{2r+4s}, on peut supposer que les entiers u, v et $w - v$ sont non nuls. Soit $r_1 = v_p(u)$ et $r_2 = v_p(w - v)$. Par (2), on a $r_1 + r_2 \geqslant 2r + 4s$.

▷ On suppose que $r_1 \geqslant r + 2s$. Alors, $v_p(u^2) \geqslant 2r + 4s$, d'où $u^2 \equiv 0\ [p^{2r+4s}]$. Par (1), on a alors

$$w + v \equiv -2(a+b)\ [p^{2r+4s}].$$

Par (3), on obtient

$$16ab = \big(-2(a+b)\big)^2 - 4(a-b)^2 \equiv (w+v)^2 - 4vw \equiv (w-v)^2\ [p^{2r+4s}].$$

Si $p > 2$, l'entier 2 est inversible modulo p^{2r+4s}, et l'on en déduit que ab est un carré modulo p^{2r+4s}. Comme $2r + 4s \geqslant r + s \geqslant v_p(ab) + 4 > v_p(ab)$, le corollaire XVII-3.5 implique que ab est un carré modulo p^k pour tout $k \geqslant 1$. Supposons que $p = 2$. L'équation précédente implique que ab est un carré modulo $2r + 4s - 4$. Puisque $r \geqslant 3$, on a $2r + 4s - 4 \geqslant r + s - 1 \geqslant v_p(ab) + 3$, et l'on conclut comme précédemment.

▷ On suppose que $r_1 < r+2s$. Alors, $r_2 > r+2s$, soit encore $r_2 \geqslant r+2s+1$. On a alors $w \equiv v\ [p^{r+2s+1}]$. Puisque $2r+4s \geqslant r+2s+1$, en utilisant (1) et (3), on obtient

$$2v + 2(a+b) \equiv u^2\ [p^{r+2s+1}], \qquad (4)$$

ainsi que

$$v^2 \equiv (a-b)^2\ [p^{r+2s+1}]. \qquad (5)$$

Écrivons $v = p^\ell v_0, p \nmid v_0$. Il existe donc $m \in \mathbb{Z}$ tel que

$$p^{2\ell} v_0^2 = p^{2s} t_0^2 + m p^{r+2s+1}.$$

En comparant les valuations p-adiques, on obtient $\ell = s$. Mais alors,

$$v_0^2 \equiv t_0^2\ [p^{r+1}].$$

Si $p > 2$, le lemme XVII-3.6 (1) implique alors que $v_0 \equiv \pm t_0\ [p^{r+1}]$. En multipliant par p^s, on obtient $v \equiv \pm(a-b)\ [p^{r+s+1}]$ puisque $\ell = s$. En utilisant (4), on obtient

$$\pm 2(a-b) + 2(a+b) \equiv u^2\ [p^{r+s+1}].$$

Autrement dit, $4a$ ou $4b$ est un carré modulo p^{r+s+1}. Comme 2 est inversible modulo p^{r+s+1}, on en déduit que a ou b est un carré modulo p^{r+s+1}. Puisque

$$r+s+1 \geqslant r+s \geqslant v(ab)+4,$$

on a aussi $r+s+1 > v_p(a)$ et $r+s+1 > v_p(b)$, et l'on conclut en utilisant le corollaire XVII-3.5.

Supposons maintenant que $p = 2$. Par le lemme XVII-3.6 (2), on a

$$v_0 \equiv \pm t_0\ [2^{r+1}] \quad \text{ou} \quad v_0 \equiv \pm t_0 + 2^r\ [2^{r+1}].$$

On a donc $v \equiv \pm(a-b)\ [2^{r+s+1}]$ ou $v \equiv \pm(a-b) + 2^{r+s}\ [2^{r+s+1}]$. On a donc $2v \equiv \pm 2(a-b)\ [2^{r+s+1}]$ dans tous les cas. En utilisant (4), on obtient

$$\pm 2(a-b) + 2(a+b) \equiv u^2\ [2^{r+s+1}].$$

On a donc $4a \equiv u^2\ [2^{r+s+1}]$ ou $4b \equiv u^2\ [2^{r+s+1}]$. En particulier, u est pair. En écrivant $u = 2u_1$, on en déduit que a ou b est un carré modulo 2^{r+s-1}. Mais $r+s-1 \geqslant v_p(ab)+3$, et par conséquent, on a $r+s-1 \geqslant v_p(a)+3$, ainsi que $r+s-1 \geqslant v_p(b)+3$. On utilise alors le corollaire XVII-3.5 pour conclure. Cela achève la démonstration. □

5. Polynômes cyclotomiques et applications

Notation. Si $n \geqslant 1$ est un entier, on note $\mathbb{U}_n$ le sous-groupe de $\mathbb{C}^\times$ constitué des racines n-ièmes de l'unité, à savoir

$$\mathbb{U}_n = \{\zeta \in \mathbb{C}^\times \mid \zeta^n = 1\}.$$

On note aussi $\mathbb{U}_n^*$ l'ensemble des éléments de $\mathbb{U}_n$ d'ordre n.

Autrement dit, on a

$$\mathbb{U}_n = \left\{ e^{\frac{2ik\pi}{n}} \mid k \in [\![0, n-1]\!] \right\} = \left\langle e^{\frac{2i\pi}{n}} \right\rangle.$$

D'après le corollaire VI-4.16, puisque $\mathbb{U}_n$ est cyclique d'ordre n, engendré par $e^{\frac{2i\pi}{n}}$, on a

$$\mathbb{U}_n^* = \left\{ e^{\frac{2ik\pi}{n}} \mid k \in [\![0, n-1]\!],\ \operatorname{pgcd}(k, n) = 1 \right\}.$$

Remarquons que, par définition de la fonction indicatrice d'Euler, on a

$$|\mathbb{U}_n^*| = \varphi(n), \quad \text{pour tout } n \geqslant 1.$$

5.1. Définition. Soit $n \geqslant 1$ un entier. Le *n-ième polynôme cyclotomique* est le polynôme de $\mathbb{C}[X]$ défini par

$$\Phi_n = \prod_{\zeta \in \mathbb{U}_n^*} (X - \zeta) \in \mathbb{C}[X].$$

C'est un polynôme unitaire de degré $\varphi(n)$.

5.2. Exemples

(1) On a clairement $\Phi_1 = X - 1$ et $\Phi_2 = X + 1$.

(2) Soit p un nombre premier. Alors, on a $\mathbb{U}_p^* = \mathbb{U}_p \setminus \{1\}$, puisque tout entier $k \in [\![1, p-1]\!]$ est premier à p. On a donc $X^p - 1 = (X-1)\Phi_p$, d'où

$$\Phi_p = X^{p-1} + X^{p-2} + \cdots + X + 1.$$

Nous allons maintenant démontrer quelques propriétés des polynômes cyclotomiques. On commence par démontrer que ce sont en fait des polynômes à coefficients entiers.

5.3. Lemme. *Pour tout $n \geqslant 1$, on a*

$$X^n - 1 = \prod_{d \mid n} \Phi_d.$$

Démonstration. Montrons que $\mathbb{U}_n$ est l'union disjointe des ensembles $\mathbb{U}_d^*$ lorsque $d \mid n$. Clairement, tout élément de $\mathbb{U}_d^*$ est aussi une racine n-ième de l'unité, et l'on a

$$\bigcup_{d \mid n} \mathbb{U}_d^* \subset \mathbb{U}_n.$$

Inversement, si $\zeta \in \mathbb{U}_n$, et si d est l'ordre de ζ dans $\mathbb{C}^\times$, on a $d \mid n$. En particulier, $\zeta^d = 1$ et $\zeta \in \mathbb{U}_d$. Mais, ζ étant d'ordre d, on a $\zeta \in \mathbb{U}_d^*$.

On a donc bien
$$\mathbb{U}_n = \bigcup_{d|n} \mathbb{U}_d^*.$$

Remarquons maintenant que les éléments de $\mathbb{U}_d^*$ sont d'ordre d dans $\mathbb{C}^\times$, donc les ensembles $\mathbb{U}_d^*$, $d \mid n$ sont deux à deux disjoints. On a alors
$$X^n - 1 = \prod_{\zeta \in \mathbb{U}_n} (X - \zeta) = \prod_{d|n} \prod_{\zeta \in \mathbb{U}_d^*} (X - \zeta) = \prod_{d|n} \Phi_d.$$

Cela achève la démonstration. □

5.4. Corollaire. *Pour tout $n \geqslant 1$, $\Phi_n \in \mathbb{Z}[X]$.*

Démonstration. Pour tout $n \geqslant 1$, on note (H_n) la propriété suivante :

(H_n) Pour tout $m \in [\![1, n]\!]$, Φ_m est un polynôme unitaire de $\mathbb{Z}[X]$.

Clairement, (H_1) est vraie puisque $\Phi_1 = X - 1 \in \mathbb{Z}[X]$ et est unitaire. Supposons maintenant que (H_n) soit vraie pour un entier $n \geqslant 1$, et montrons (H_{n+1}). Il reste à démontrer que $\Phi_{n+1} \in \mathbb{Z}[X]$ et est unitaire. Par le lemme 5.3, on a
$$X^{n+1} - 1 = \Phi_{n+1} \prod_{d|n+1,\, d<n+1} \Phi_d.$$

Comme les diviseurs stricts de $n+1$ sont inférieurs ou égaux à n, l'hypothèse de récurrence montre que le polynôme $Q = \prod_{d|n+1, d<n+1} \Phi_d$ est un polynôme unitaire de $\mathbb{Z}[X]$. On peut alors effectuer la division de $X^{n+1} - 1$ par Q dans $\mathbb{Z}[X]$ d'après le théorème XV-6.10. Mais le quotient et le reste de cette division, qui sont des polynômes de $\mathbb{Z}[X]$, sont aussi le quotient et le reste de la division dans $\mathbb{C}[X]$. Par unicité, ce quotient est Φ_{n+1}, on a donc que $\Phi_{n+1} \in \mathbb{Z}[X]$ et est unitaire car Q et $X^{n+1} - 1$ le sont. Cela achève la récurrence, ainsi que la démonstration. □

La formule d'inversion de Moëbius, appliquée à la relation $X^n - 1 = \prod_{d|n} \Phi_d$ dans le groupe multiplicatif $\mathbb{Q}(X)^\times$, donne également le résultat suivant.

5.5. Corollaire. *Pour tout $n \geqslant 1$, on a*
$$\Phi_n = \prod_{d|n} \left(X^{\frac{n}{d}} - 1\right)^{\mu(d)},$$
où μ est la fonction de Möbius.

Cette formule permet en théorie de calculer les coefficients de tous les polynômes cyclotomiques de manière explicite. Néanmoins, nous allons main-

tenant établir des relations entre les polynômes cyclotomiques qui permettront de réduire les calculs. On commence par un lemme.

5.6. Lemme. *Pour tous entiers $n, m \geqslant 1$ distincts, les polynômes Φ_n et Φ_m sont premiers entre eux dans $\mathbb{Q}[X]$ et $\mathbb{C}[X]$.*

Démonstration. Remarquons que pour tout entier $k \geqslant 1$, les éléments de $\mathbb{U}_k^*$ sont des éléments d'ordre k de $\mathbb{C}^\times$. En particulier, si $n \neq m$, alors Φ_n et Φ_m n'ont pas de racines communes dans $\mathbb{C}$. A fortiori, ils n'ont donc pas de diviseurs communs non constants, et sont donc premiers entre eux dans l'anneau $\mathbb{C}[X]$. Un diviseur commun dans $\mathbb{Q}[X]$ étant aussi un diviseur commun dans $\mathbb{C}[X]$, ils sont également premiers entre eux dans $\mathbb{Q}[X]$. □

5.7. Proposition. *Soit $n \geqslant 1$ et soit p un nombre premier.*

(1) *Si $p \nmid n$, on a $\Phi_n(X)\Phi_{np}(X) = \Phi_n(X^p)$.*

(2) *Si $p \mid n$, on a $\Phi_{np}(X) = \Phi_n(X^p)$.*

(3) *Si $n \geqslant 3$ est impair, on a $\Phi_{2n}(X) = \Phi_n(-X)$.*

Démonstration. Montrons (1). Par le corollaire XVII-2.3 (1), puisque $p \nmid n$, on a

$$\varphi(np) = \varphi(n)\varphi(p) = (p-1)\varphi(n).$$

Mais alors, on a

$$\deg(\Phi_n\Phi_{np}) = \varphi(n) + \varphi(np) = p\varphi(n) = \deg\big(\Phi_n(X^p)\big).$$

Les polynômes de l'égalité étant tous deux unitaires et de même degré, il suffit par exemple de démontrer que $\Phi_n(X)\Phi_{np}(X) \mid \Phi_n(X^p)$.

Comme les polynômes Φ_n et Φ_{np} sont premiers entre eux par le lemme précédent, il suffit donc de vérifier que $\Phi_n \mid \Phi_n(X^p)$ et $\Phi_{np} \mid \Phi_n(X^p)$. Or, si ζ est une racine n-ième de l'unité d'ordre n, il en est de même de ζ^p, puisque p est premier à n. Autrement dit, si ζ est une racine complexe de Φ_n, c'est aussi une racine complexe de $\Phi_n(X^p)$, d'où $\Phi_n \mid \Phi_n(X^p)$. Enfin, si ζ' est une racine np-ième de l'unité d'ordre np, alors ζ'^p est une racine n-ième de l'unité d'ordre n, d'où $\Phi_{np} \mid \Phi_n(X^p)$. Cela établit (1).

Si $p \mid n$, écrivons $n = mp^r$, avec $r \geqslant 1$ et $p \nmid m$. On a alors

$$\varphi(np) = \varphi(mp^{r+1}) = \varphi(p^{r+1})\varphi(m) = p^r(p-1)\varphi(m).$$

On a ainsi

$$\varphi(np) = p\big(p^{r-1}(p-1)\varphi(m)\big) = p\varphi(n).$$

Comme précédemment, il suffit donc de montrer que $\Phi_{np} \mid \Phi_n(X^p)$, ce qui se vérifie comme on vient de le faire plus haut.

Il reste donc à établir le point (3). Supposons $n \geqslant 3$ impair. Si ζ est une racine n-ième de l'unité d'ordre n, puisque n est impair, on voit facilement

que $-\zeta$ est d'ordre $2n$ dans $\mathbb{C}^\times$. On a donc $-\zeta \in \mathbb{U}_{2n}^*$. Par ailleurs, on a

$$|\mathbb{U}_{2n}^*| = \varphi(2n) = \varphi(2)\varphi(n) = \varphi(n) = |-\mathbb{U}_n^*|.$$

Ainsi, $\mathbb{U}_{2n}^* = -\mathbb{U}_n^*$, et par conséquent

$$\Phi_{2n} = \prod_{\zeta \in \mathbb{U}_n^*} (X + \zeta) = (-1)^{\varphi(n)} \Phi_n(-X).$$

Comme n est impair, la formule explicite donnant le calcul de $\varphi(n)$ (cf. corollaire XVII-2.3 (3)) montre que $\varphi(n)$ est pair (car $n \geqslant 3$). Cela achève la démonstration. □

5.8. Remarque. Une application répétée de la proposition précédente permet de démontrer par exemple les formules suivantes.

(1) Pour tout p premier, tout n premier à p, et tout $r \geqslant 1$, on a

$$\Phi_{np^r} = \frac{\Phi_n(X^{p^r})}{\Phi_n(X^{p^{r-1}})}.$$

En particulier, on a

$$\Phi_{p^r} = \Phi_p(X^{p^{r-1}}) = X^{(p-1)p^{r-1}} + X^{(p-2)p^{r-1}} + \cdots + X^{p^{r-1}} + 1.$$

(2) Pour tout $n \geqslant 3$ impair, et tout $r \geqslant 1$, on a

$$\Phi_{2^r n} = \Phi_n(-X^{2^{r-1}}).$$

Ces formules permettent de calculer Φ_n de proche en proche en connaissant la décomposition en facteurs premiers de n. La démonstration est laissée au lecteur à titre d'exercice.

Nous en arrivons au premier résultat important de ce paragraphe.

5.9. Théorème. *Pour tout $n \geqslant 1$, le polynôme Φ_n est irréductible dans les anneaux $\mathbb{Z}[X]$ et $\mathbb{Q}[X]$.*

Démonstration. La démonstration qui suit est due à Landau.

Par la remarque 3.11, il suffit de démontrer l'irréductibilité de Φ_n dans l'anneau $\mathbb{Z}[X]$, puisque $\mathbb{Z}$ est factoriel.

Puisque Φ_n est unitaire, le coefficient dominant d'un facteur irréductible P de Φ_n dans $\mathbb{Z}[X]$ est inversible dans $\mathbb{Z}$, donc égal à ± 1. Quitte à remplacer P par $-P$, on peut supposer P unitaire. Il est en particulier non constant.

Pour tout $k \geqslant 0$, on note R_k le reste de la division de $P(X^k)$ par P. En particulier, $\deg(R_k) < \deg(P)$ pour tout $k \geqslant 0$. Remarquons que R_k est aussi le reste de la division de $P(X^k) - P(X)^k$ par P. Enfin, si $\zeta \in \mathbb{U}_n^*$ est une racine de P, on a $P(\zeta^k) = R_k(\zeta)$, puisque $P(X^k) - R_k$ est un multiple de P.

On commence par montrer que si l'on écrit $k = qn + s$, avec $s \in [\![0, n-1]\!]$, alors $R_k = R_s$.

Comme $\zeta^n = 1$, on a $\zeta^k = \zeta^s$, et par suite

$$R_k(\zeta) = P(\zeta^k) = P(\zeta^s) = R_s(\zeta).$$

Ainsi, $R_k - R_s$ s'annule en toute racine de P. Or, Φ_n étant à racines simples, il en est de même de P, et on en conclut que $P \mid R_k - R_s$. Or, on a

$$\deg(R_k - R_s) < \deg(P),$$

et donc nécessairement $R_k - R_s = 0$, soit $R_k = R_s$.

Avant de continuer, remarquons que pour tout nombre premier p, et tout polynôme $Q \in \mathbb{F}_p[X]$, on a $Q(X^p) = Q(X)^p \in \mathbb{F}_p[X]$.

En effet, puisque $\mathbb{F}_p[X]$ est de caractéristique p, si $Q = \overline{a}_d X^d + \cdots + \overline{a}_1 X + \overline{a}_0$, on a

$$Q(X)^p = \overline{a}_d^p X^{dp} + \cdots + \overline{a}_1^p X^p + \overline{a}_0^p,$$

et le petit théorème de Fermat donne alors

$$Q(X)^p = \overline{a}_d X^{dp} + \cdots + \overline{a}_1 X^p + \overline{a}_0 = Q(X^p).$$

Dans $\mathbb{Z}[X]$, cette identité se retraduit de la manière suivante : pour tout polynôme $S \in \mathbb{Z}[X]$, on a $S(X^p) - S(X)^p \in p\mathbb{Z}[X]$.

Cela étant dit, soit $\ell \geqslant 1$ un entier majorant les valeurs absolues de tous les coefficients des polynômes $R_0, \ldots, R_{n-1}$. Fixons un nombre premier $p > \ell$. Alors, $P(X^p) - P(X)^p \in p\mathbb{Z}[X]$ d'après ce qui précède. Mais alors, le reste de la division de $P(X^p) - P(X)^p$ par P est aussi dans $p\mathbb{Z}[X]$. Autrement dit, $R_p \in p\mathbb{Z}[X]$. En effet, si l'on écrit $P(X^p) - P(X)^p = pS$, $S \in \mathbb{Z}[X]$, le reste de la division de $P(X^p) - P(X)^p$ par P est égal à p fois le reste de la division de S par P, par unicité du reste.

Or, d'après le premier point, $R_p = R_s$ pour un certain entier $s \in [\![0, n-1]\!]$. Si R_p était non nul, il aurait un coefficient de valeur absolue $\geqslant p$, ce qui contredit le choix de p. Par conséquent, $R_p = 0$, d'où $P \mid (P(X^p) - P(X)^p)$, ou encore $P \mid P(X^p)$ pour tout $p > \ell$. Il en découle que, pour tout nombre premier $p > \ell$, et pour toute racine $\zeta \in \mathbb{U}_n^*$ de P, ζ^p est aussi une racine de P.

Une récurrence facile montre que pour tout entier $m \geqslant 1$ dont tous les facteurs premiers sont $> \ell$, et pour toute racine $\zeta \in \mathbb{U}_n^*$ de P, ζ^m est une racine de P. Notons qu'a priori, rien ne garantit que ζ^m soit d'ordre n.

Fixons maintenant $\zeta_0 \in \mathbb{U}_n^*$ une racine de P. Soit $k \in [\![1, n]\!]$ un entier premier à n. Soit q le produit de tous les nombres premiers $p \leqslant \ell$ ne divisant pas k, et soit $m = nq + k$. Soit p un nombre premier $\leqslant \ell$. Si $p \mid k$, alors $p \nmid q$ par définition, et $p \nmid n$, car k et n sont premiers entre eux. Ainsi,

on a $p \nmid nq$, car p est premier, et donc $p \nmid m$. Si $p \nmid k$, alors on a $p \mid q$, et donc $p \mid nq$. Par conséquent, $p \nmid m$ dans ce cas également. Autrement dit, tous les diviseurs premiers de m sont $> \ell$.

D'après le premier point, ζ_0^m est une racine de P. Comme $m = nq + k$, on en déduit que ζ_0^k est une racine de P, et cela pour tout entier $k \in [\![1, n]\!]$ premier à n. Or, comme ζ_0 est d'ordre n, les divers éléments

$$\zeta_0^k, \quad k \in [\![1, n]\!], \ \operatorname{pgcd}(k, n) = 1$$

sont $\varphi(n)$ éléments de $\mathbb{U}_n^*$ deux à deux distincts, et décrivent donc $\mathbb{U}_n^*$ tout entier. Il s'ensuit que $\Phi_n \mid P$, d'où $P = \Phi_n$, car P et Φ_n sont tous deux unitaires. On en conclut que Φ_n est irréductible dans $\mathbb{Z}[X]$, ce qui achève la démonstration. □

5.10. Remarque. On obtient en particulier que pour tout p premier, le polynôme

$$\Phi_p = X^p + X^{p-1} + \cdots + X + 1$$

est irréductible dans $\mathbb{Z}[X]$ et $\mathbb{Q}[X]$. On peut retrouver ce fait plus simplement comme suit. Comme précédemment, il suffit d'établir l'irréductibilité dans $\mathbb{Z}[X]$.

D'après le lemme 1.6, il suffit de démontrer que $\Phi_p(X+1)$ est irréductible dans $\mathbb{Z}[X]$. L'identité $(X-1)\Phi_p = X^p - 1$ donne

$$X\Phi_p(X+1) = (X+1)^p - 1,$$

d'où

$$\Phi_p(X+1) = X^{p-1} + \binom{p}{1} X^{p-2} + \cdots + \binom{p}{p-2} X + \binom{p}{p-1}.$$

Tous ces coefficients binomiaux sont divisibles par p, et $\binom{p}{p-1} = p$ n'est pas divisible par p^2. Ainsi, $\Phi_p(X+1)$ est p-Eisenstein, donc irréductible dans l'anneau $\mathbb{Z}[X]$.

Nous allons maintenant donner deux applications des polynômes cyclotomiques. Tout d'abord, nous allons démontrer un cas particulier du théorème de la progression arithmétique de Dirichlet. Ce théorème s'énonce comme suit : pour tout entier $n \geqslant 2$, et tout $a \in \mathbb{Z}$ premier à n, il existe une infinité de nombres premiers p congrus à a modulo n.

La démonstration du cas général est assez délicate, et requiert de l'analyse complexe (cf. [43]), mais le cas $a = 1$ se démontre de manière simple et purement algébrique grâce aux polynômes cyclotomiques.

5.11. Théorème (de la progression arithmétique de Dirichlet – version faible). *Soit $n \geqslant 2$ un entier. Alors, il existe une infinité de nombres premiers congrus à 1 modulo n.*

Démonstration. Commençons par une remarque préliminaire. Soit $N \geqslant 2$ un entier. On suppose que l'on dispose d'un entier $a \in \mathbb{Z}$ et d'un nombre premier p tels que $p \mid \Phi_N(a)$ et $p \nmid \Phi_d(a)$ pour tout diviseur strict d de N. Alors, $p \equiv 1\ [N]$.

En effet, puisque $\Phi_N \mid X^N - 1$, on a également $\Phi_N(a) \mid a^N - 1$, et par suite $p \mid a^N - 1$, soit $\overline{a}^N = \overline{1} \in \mathbb{F}_p$. Si maintenant d est un diviseur strict de N, alors l'identité $X^d - 1 = \prod_{d' \mid d} \Phi_{d'}$ (cf. lemme 5.3) implique en particulier que

$$a^d - 1 = \prod_{d' \mid d} \Phi_{d'}(a).$$

Par hypothèse, $\Phi_{d'}(a)$ n'est pas congru à 0 modulo p, et comme $\mathbb{F}_p$ est intègre, on en déduit que $p \nmid a^d - 1$. Autrement dit, $\overline{a}^d \neq \overline{1} \in \mathbb{F}_p$ pour tout diviseur strict d de N. Il s'ensuit que l'ordre de $\overline{a}$ dans le groupe $\mathbb{F}_p^\times$ est exactement N. Comme $\mathbb{F}_p^\times$ est d'ordre $p-1$, le théorème de Lagrange nous assure que $N \mid p - 1$, d'où $p \equiv 1\ [N]$.

Cela étant dit, montrons maintenant qu'étant fixé un entier $N \geqslant 2$, on peut toujours trouver un entier $a \in \mathbb{Z}$ et un nombre premier p vérifiant les conditions précédentes. Soit P le produit des Φ_d lorsque d parcourt les diviseurs stricts de N. Par le lemme 5.3, on a donc

$$X^N - 1 = \Phi_N P,$$

et le lemme 5.6 assure que Φ_N et P sont premiers entre eux dans l'anneau $\mathbb{Q}[X]$. Il existe donc $U, V \in \mathbb{Q}[X]$ tels que

$$U\Phi_N + VP = 1.$$

Si $r \in \mathbb{Z}$ est un entier tel que $rU, rV \in \mathbb{Z}[X]$, on a donc

$$(rU)\Phi_N + (rV)P = r.$$

Puisque Φ_N est non constant, les entiers $b \in \mathbb{Z}$ tels que $\Phi_N(br) = \pm 1$ sont en nombre fini (sinon, un des deux polynômes $\Phi_N \pm 1$ aurait une infinité de racines, donc serait nul, et Φ_N serait constant). Il existe donc $b \in \mathbb{Z}$ tel que $|\Phi_N(br)| \geqslant 2$. Posons $a = br$. Alors, $\Phi_N(a)$ possède donc au moins un diviseur premier p. De plus, en multipliant la relation précédente par b et en évaluant en a, on obtient l'existence de deux entiers $u, v \in \mathbb{Z}$ tels que

$$u\Phi_N(a) + vP(a) = a.$$

Puisque $p \mid \Phi_N(a)$, on a $p \mid a^N - 1$, et ainsi $p \nmid a$. Mais alors, la relation ci-dessus montre que $p \nmid P(a)$, ce qui revient à dire que $p \nmid \Phi_d(a)$ pour tout diviseur strict d de N par définition de P et par le lemme d'Euclide.

Nous avons maintenant tous les ingrédients pour construire une infinité de nombres premiers p congrus à 1 modulo n, pour un entier $n \geqslant 2$ fixé.

Soit $n \geqslant 2$ un entier. Par ce qui précède, on peut trouver un entier $a \in \mathbb{Z}$ et un nombre premier p_1 tel que $p_1 \mid \Phi_n(a)$ et $p_1 \nmid \Phi_d(a)$ pour tout diviseur strict d de n. D'après le premier point, $p_1 \equiv 1\ [n]$.

Supposons maintenant que l'on ait trouvé des nombres premiers $p_1, \ldots, p_r$ congrus à 1 modulo n deux à deux distincts. Posons $N = np_1 \cdots p_r$. Comme précédemment, on peut construire un nombre premier p_{r+1} congru à 1 modulo N. En particulier, $p_{r+1} \equiv 1\ [n]$, ainsi que

$$p_{r+1} \equiv 1\ [p_i], \quad \text{pour tout } i \in [\![1, r]\!].$$

Ces dernières congruences nous assurent en particulier que p_{r+1} est distinct de $p_1, \ldots, p_r$. Cela achève la démonstration. □

Nous allons conclure ce paragraphe en démontrant un théorème dû à Wedderburn, à savoir que tout anneau à division fini est commutatif[3].

Nous allons tout d'abord démontrer un lemme calculatoire.

5.12. Lemme. *Soient n, d, $q \geqslant 1$ des entiers, et soit r le reste de la division euclidienne de n par d. On suppose que $q \geqslant 2$. Alors, $q^r - 1$ est le reste de la division euclidienne de $q^n - 1$ par $q^d - 1$. En particulier, $q^d - 1 \mid q^n - 1$ si, et seulement si, $d \mid n$.*
Dans ce cas, on a $\Phi_n(q) \mid \dfrac{q^n - 1}{q^d - 1}$ pour tout diviseur strict d de n.

Démonstration. Écrivons $n = md + r$, avec $m \in \mathbb{Z}$ et $r \in [\![0, d-1]\!]$. Remarquons que $m \geqslant 0$ puisque $n \geqslant 0$ et $d \geqslant 0$. On a alors

$$q^n - 1 = (q^d)^m q^r - 1 = \big((q^d - 1) + 1\big)^m q^r - 1,$$

d'où $q^n - 1 \equiv q^r - 1\ [q^d - 1]$. Il existe donc $k \in \mathbb{Z}$ tel que

$$q^n - 1 = k(q^d - 1) + q^r - 1.$$

Remarquons que $q^d - 1$ est non nul puisque $q \geqslant 2$ et $d \geqslant 1$. De plus, puisque l'on a $q \geqslant 2$ et $r \in [\![0, d-1]\!]$, il vient

$$0 \leqslant q^r - 1 < q^d - 1.$$

Ainsi, $q^r - 1$ est le reste de la division euclidienne de $q^n - 1$ par $q^d - 1$. Puisque $q \geqslant 2$, ce reste est nul si, et seulement si, $r = 0$, ce qui équivaut à dire que $d \mid n$.

3. Les auteurs autorisant un corps à ne pas être nécessairement commutatif énonceraient cela de manière plus conventionnelle comme suit : tout corps fini est commutatif.

Montrons le dernier point. Si $d \mid n$, le lemme 5.3 montre que l'on a

$$X^n - 1 = (X^d - 1) \prod_{d' \mid n, d' \nmid d} \Phi_{d'}.$$

Si $d \neq n$, en particulier $n \nmid d$ et la relation précédente montre l'existence d'un polynôme $Q \in \mathbb{Z}[X]$ tel que

$$X^n - 1 = (X^d - 1) Q \Phi_n.$$

En spécialisant en q, on en déduit que $\Phi_d(q) \mid \dfrac{q^n - 1}{q^d - 1}$. Cela achève la démonstration. □

On peut maintenant démontrer le théorème de Wedderburn.

5.13. Théorème (Wedderburn). *Tout anneau à division fini est commutatif.*

Démonstration. Soit D un anneau à division fini, et soit

$$K = Z(D) = \{z \in D \mid zd = dz, \text{ pour tout } d \in D\}.$$

On vérifie que K est un sous-anneau à division de D, qui est commutatif. C'est donc un corps. Comme D est fini, K aussi. Soit $q = |K|$. Puisque K est un corps, on a $q \geqslant 2$.

Remarquons que pour tout sous-anneau C de D contenant K, la loi

$$\begin{aligned} K \times C &\longrightarrow C \\ (z, c) &\longmapsto zc \end{aligned}$$

induit sur le groupe $(C, +)$ une structure de K-espace vectoriel, nécessairement de dimension finie puisque D est fini. Si $m = \dim_K(C)$, on a donc un isomorphisme de K-espaces vectoriels $C \simeq K^m$, et ainsi $|C| = q^m$. Remarquons que $m \geqslant 1$ puisque C n'est pas réduit à 0. En particulier, il existe un entier $n \geqslant 1$ tel que $|D| = q^n$.

Faisons maintenant agir $D^\times$ sur lui-même par conjugaison. Soit ω une orbite de $D^\times$ sous cette action, et soit $d \in \omega$. On a donc $\omega = \mathcal{O}_d$. Or, on a

$$\mathrm{Stab}_{D^\times}(d) = \{z \in D^\times \mid zdz^{-1} = z\} = \{z \in D^\times \mid zd = dz\} = C(d) \setminus \{0\},$$

où $C(d) = \{z \in D \mid zd = dz\}$. Comme $C(d)$ est un sous-anneau de D contenant K, il existe un entier $n_\omega \geqslant 1$ tel que $|C(d)| = q^{n_\omega}$.

Par conséquent,

$$|\omega| = |\mathcal{O}_d| = \frac{|D^\times|}{|\mathrm{Stab}_{D^\times}(d)|} = \frac{q^n - 1}{q^{n_\omega} - 1}.$$

En particulier, $q^{n_\omega} - 1 \mid q^n - 1$. Le lemme précédent assure alors que $n_\omega \mid n$.

Remarquons aussi que $n_\omega = n$ si, et seulement si, $|\omega| = |\mathcal{O}_d| = 1$. Cela revient à dire que d est un point fixe, ou encore que $d \in K^\times$. Il y a donc exactement $q-1$ orbites réduites à un élément.

Comme D est fini, l'ensemble des orbites Ω de $D^\times$ sous l'action de conjugaison est fini. Soit Ω' l'ensemble des orbites non réduites à un singleton. Comme $D^\times$ est l'union disjointe de ces diverses orbites, on en déduit

$$q^n - 1 = q - 1 + \sum_{\omega \in \Omega'} \frac{q^n - 1}{q^{n_\omega} - 1},$$

où les entiers n_ω sont des diviseurs stricts de n.

D'après le lemme précédent, chaque terme de la somme indexée par Ω' est divisible par $\Phi_n(q)$. C'est le cas même si Ω' est vide, puisque dans ce cas cette somme est nulle. Comme $q^n - 1 = \prod_{d|n} \Phi_d(q)$, on a $\Phi_n(q) \mid q^n - 1$. On en déduit finalement que $\Phi_n(q) \mid q - 1$. En comparant les valeurs absolues, on en déduit

$$|\Phi_n(q)| \leqslant q - 1.$$

Or, pour tout $\zeta \in \mathbb{U}_n^*$, on a

$$|q - \zeta| \geqslant \big||q| - |\zeta|\big|,$$

soit $|q - \zeta| \geqslant q - 1 \geqslant 1$. Soit $\zeta_n = e^{\frac{2i\pi}{n}}$. On a alors

$$q - 1 \geqslant |\Phi_n(q)| = \prod_{\zeta \in \mathbb{U}_n^*} |q - \zeta| \geqslant |q - \zeta_n| \geqslant q - 1.$$

On en déduit alors que $|q - \zeta_n| = q - 1$, ce qui se récrit

$$\big(q - \cos\big(\tfrac{2\pi}{n}\big)\big)^2 + \sin\big(\tfrac{2\pi}{n}\big)^2 = (q-1)^2.$$

En développant les carrés, on obtient $2q\cos(\frac{2\pi}{n}) = 2q$, soit $\cos(\frac{2\pi}{n}) = 1$. Ainsi, $n = 1$. Autrement dit, $\dim_K(D) = 1$, ou encore $D = K$ (puisque l'on a $K \subset D$). Par conséquent, D est commutatif. □

6. Polynômes irréductibles sur $\mathbb{F}_p$, corps finis

Le théorème précédent nous assure qu'un anneau à division fini est commutatif. Nous allons nous intéresser maintenant à l'existence et aux propriétés des corps finis. Commençons par quelques résultat simples.

6.1. Lemme. *Soit K un corps fini. Alors, $|K|$ est une puissance non triviale d'un nombre premier p. De plus,* $\mathrm{car}(K) = p$.

Démonstration. Puisqu'un corps est intègre, sa caractéristique c est soit nulle, soit un nombre premier p d'après le corollaire XV-3.10. Comme K

contient un sous-anneau à isomorphe $\mathbb{Z}/c\mathbb{Z}$ (toujours d'après cette même proposition), le premier cas est impossible puisque K est fini. Ainsi, K est de caractéristique p. La loi externe

$$\begin{aligned} \mathbb{F}_p \times K &\longrightarrow K \\ (\overline{m}, x) &\longmapsto m{\cdot}x \end{aligned}$$

est alors bien définie[4] et induit sur $(K,+)$ une structure de $\mathbb{F}_p$-espace vectoriel, nécessairement de dimension finie car K est un corps fini. Si l'on pose $n = \dim_{\mathbb{F}_p}(K)$, alors $n \geqslant 1$ puisqu'un corps est un anneau non trivial. On obtient alors un isomorphisme de $\mathbb{F}_p$- espaces vectoriels $K \simeq \mathbb{F}_p^n$, et en particulier $|K| = p^n$. Cela achève la démonstration. □

Le nombre d'éléments d'un corps fini est donc nécessairement une puissance non triviale d'un nombre premier. On peut donc légitimement se poser les questions suivantes : soit q une puissance non triviale d'un nombre premier.

(1) Existe-t-il au moins un corps à q éléments ?

De plus, s'il est clair qu'il n'existe à isomorphisme près qu'un nombre fini de corps à q éléments, on peut se poser la question de leur nombre.

(2) Combien de corps à q éléments y a-t-il, à isomorphisme près ?

Le but de ce qui suit est de répondre à ces deux questions. Remarquons que, si l'on dispose d'un polynôme irréductible $f \in \mathbb{F}_p[X]$ de degré $n \geqslant 1$, alors on peut construire un corps à p^n éléments très facilement, comme le montre le lemme suivant.

6.2. Lemme. *Soit $f \in \mathbb{F}_p[X]$ un polynôme irréductible de degré $n \geqslant 1$. Alors, $\mathbb{F}_p[X]/(f)$ est un corps. De plus, c'est aussi un $\mathbb{F}_p$-espace vectoriel de base $\left(\overline{1}, \overline{X}, \ldots, \overline{X}^{n-1}\right)$. En particulier, il a p^n éléments.*

Démonstration. Tout découle de la proposition 1.9, page 649. □

Le lemme précédent nous conduit à nous poser la question suivante : étant donné un nombre premier p, existe-t-il des polynômes irréductibles à coefficients dans $\mathbb{F}_p$ de tout degré ?

Nous établirons que la réponse à cette question est positive, ce qui assurera l'existence d'un corps fini à p^n éléments pour tout p premier et tout $n \geqslant 1$.

Nous aurons besoin de plusieurs résultats intermédiaires.

4. Comme le lecteur le vérifiera sans peine...

6.3. Lemme. *Soit K un corps fini à q éléments. Alors, pour tout $x \in K$, on a $x^q - x = 0$.*

Démonstration. Puisque K est un corps, $K^\times$ est un groupe d'ordre $q-1$. D'après le théorème de Lagrange, on a

$$x^{q-1} = 1, \quad \text{pour tout } x \in K^\times.$$

Ainsi, $x^q = x$ pour tout $x \in K^\times$. Cela étant vrai aussi pour $x = 0$, on a le résultat voulu. □

Nous allons maintenant nous intéresser à la factorisation de $X^{p^n} - X$ en produit de facteurs irréductibles dans $\mathbb{F}_p[X]$.

6.4. Lemme. *Soit $f \in \mathbb{F}_p[X]$ un polynôme irréductible de degré $d \geqslant 1$. Alors, $f \mid X^{p^d} - X$.*

Démonstration. Soit $K = \mathbb{F}_p[X]/(f)$. D'après le lemme 6.2, K est un corps fini à p^d éléments. Le lemme précédent donne alors

$$\overline{X^{p^d} - X} = \overline{X}^{p^d} - \overline{X} = \overline{0},$$

d'où $f \mid X^{p^d} - X$. □

6.5. Lemme. *Soit $f \in \mathbb{F}_p[X]$ un polynôme irréductible de degré $d \geqslant 1$, et soit $n \geqslant 1$. Alors, $f \mid X^{p^n} - X$ si, et seulement si, $d \mid n$.*

Démonstration. Supposons que $d \mid n$.

Par le lemme 5.12, $p^d - 1 \mid p^n - 1$. Il existe donc un entier $m \geqslant 1$ tel que $p^n - 1 = (p^d - 1)m$. Mais alors, on a

$$X^{p^n-1} - 1 = (X^{p^d-1})^m - 1 = (X^{p^d-1} - 1)\big(1 + X^{p^d-1} + \cdots + (X^{p^d-1})^{m-1}\big).$$

Par conséquent, $X^{p^d-1} - 1 \mid X^{p^n-1} - 1$, et ainsi $X^{p^d} - X \mid X^{p^n} - X$. On en conclut alors que $f \mid X^{p^n} - X$ par le lemme 6.4.

Inversement, supposons que $f \mid X^{p^n} - X$, et posons $K = \mathbb{F}_p[X]/(f)$. C'est un corps à p^d éléments, qui est engendré en tant qu'anneau par $\overline{X}$.

Montrons l'égalité

$$K = \{x \in K \mid x^{p^n} = x\}.$$

On vérifie facilement que $K' = \{x \in K \mid x^{p^n} = x\}$ est un sous-anneau de K en utilisant le lemme XV-1.18. Par hypothèse sur f, on a $\overline{X}^{p^n} = \overline{X}$, soit encore $\overline{X} \in K'$. Par conséquent, $K \subset K'$. L'autre inclusion étant évidente, on a $K = K'$. En particulier, on a

$$K^\times = \{x \in K \mid x^{p^n-1} = \overline{1}\}.$$

Or, le groupe $K^\times$ étant cyclique par le théorème XII-2.8, il existe $\alpha \in K^\times$ d'ordre $|K^\times| = p^d - 1$. Comme $\alpha^{p^n-1} = \overline{1}$, on en déduit que $p^d - 1 \mid p^n - 1$, d'où $d \mid n$ par le lemme 5.12. Cela achève la démonstration. □

6.6. Corollaire. *Soit p un nombre premie et soit $n \geqslant 1$ un entier. Alors, le polynôme $X^{p^n} - X \in \mathbb{F}_p[X]$ est le produit de tous les polynômes irréductibles unitaires de $\mathbb{F}_p[X]$ de degré divisant n.*

Démonstration. Notons $\mathscr{U}$ l'ensemble des polynômes irréductibles unitaires de $\mathbb{F}_p$ de degré divisant n. D'après le lemme précédent, les facteurs irréductibles unitaires de $X^{p^n} - X$ sont exactement les éléments de $\mathscr{U}$. Montrons maintenant que pour tout $f \in \mathbb{F}_p[X] \in \mathscr{U}$, $f^2 \nmid X^{p^n} - X$. Dans le cas contraire, on pourrait écrire

$$X^{p^n} - X = f^2 g,$$

avec $g \in \mathbb{F}_p[X]$ et $f \in \mathscr{U}$. En appliquant la dérivée formelle, on obtiendrait alors $(p^n \cdot \overline{1})X^{p^n-1} - \overline{1} = f(2f'g + fg')$, soit

$$-\overline{1} = f(2f'g + fg') \in \mathbb{F}_p[X],$$

puisque $\mathbb{F}_p[X]$ est de caractéristique p. En particulier, f serait constant, d'où une contradiction. Finalement, $X^{p^n} - X$ et $\prod_{f \in \mathscr{U}} f$ diffèrent par une constante multiplicative, égale à 1 puisque les deux polynômes sont unitaires. Cela achève la démonstration. □

On peut maintenant démontrer l'existence de polynômes irréductibles unitaires de $\mathbb{F}_p[X]$ de degré fixé.

6.7. Proposition. *Soit p un nombre premier. Pour tout entier $d \geqslant 1$, on note $I(d,p)$ le nombre de polynômes irréductibles unitaires de $\mathbb{F}_p[X]$ de degré d. Alors, pour tout entier $n \geqslant 1$, on a*

(1) $p^n = \sum_{d \mid n} dI(d,p)$;

(2) $I(n,p) = \frac{1}{n} \sum_{d \mid n} \mu(\frac{n}{d}) p^d$;

(3) $\frac{p^n - p^{m+1}}{n} < I(n,p) \leqslant \frac{p^n}{n}$, *où m est la partie entière de* $\frac{n}{2}$;

(4) *il existe un polynôme irréductible unitaire de $\mathbb{F}_p[X]$ de degré n.*

Démonstration. En utilisant le corollaire précédent et en comparant les degrés, on obtient (1). Le point (2) découle alors de la formule d'inversion de Möbius, appliquée au groupe abélien $(\mathbb{Z}, +)$.

Comme $nI(n,p) \leqslant \sum_{d|n} dI(d,p)$, le point (1) donne la majoration

$$I(n,p) \leqslant \frac{p^n}{n}.$$

En appliquant cette majoration à tous les diviseurs stricts de n, on obtient

$$p^n - nI(n,p) \leqslant \sum_{d|n,\, d\neq n} p^d.$$

Remarquons qu'un diviseur strict de n est nécessairement $\leqslant m$, où m est la partie entière de $\frac{n}{2}$. Mais alors, on a

$$p^n - nI(n,p) \leqslant \sum_{r=1}^{m} p^r,$$

d'où

$$p^n - nI(n,p) \leqslant \frac{p^{m+1}-1}{p-1} < p^{m+1}.$$

On obtient ainsi la minoration annoncée, d'où le point (3). Enfin, par définition, on a soit $n = 2m$, auquel cas $m \geqslant 1$, soit $n = 2m+1$, auquel cas $m \geqslant 0$. Dans chacune des éventualités, on a $n \geqslant m+1$, et ainsi $p^n - p^{m+1} \geqslant 0$. Le point (3) montre alors que $I(n,p) > 0$. Comme $I(n,p)$ est un entier, on en déduit que $I(n,p) \geqslant 1$, d'où (4). Cela achève la démonstration. □

6.8. Exemple. Puisque tout polynôme de degré 1 à coefficients dans un corps est irréductible, on a $I(1,p) = p$.

Si ℓ est un nombre premier, on a $p^\ell = I(1,p) + \ell I(\ell,p)$, et ainsi

$$I(\ell,p) = \frac{p^\ell - p}{\ell}, \quad \text{pour tout nombre premier } \ell.$$

Par exemple, $I(2,2) = 1$ et $I(3,2) = 2$. Il est d'ailleurs aisé de trouver les polynômes irréductibles unitaires sur $\mathbb{F}_2[X]$ de degré 2 ou 3. On pourrait bien entendu factoriser $X^4 - X$ et $X^8 - X$, mais il est tout aussi simple de remarquer qu'un tel polynôme est irréductible si, et seulement si, il n'a pas de racines dans $\mathbb{F}_2$, ce qui revient à dire que $f(\overline{0}) = \overline{1}$ et f possède un nombre impair de coefficients non nuls. On trouve alors les polynômes suivants

$$X^2 + X + \overline{1}, \quad X^3 + X + \overline{1}, \quad X^3 + X^2 + \overline{1} \in \mathbb{F}_2[X].$$

On vérifie d'ailleurs par le calcul que l'on a bien

$$X^4 - X = X = X(X+\overline{1})(X^2 + X + \overline{1}) \in \mathbb{F}_2[X],$$

ainsi que

$$X^8 - X = X(X+\overline{1})(X^3 + X + \overline{1})(X^3 + X^2 + \overline{1}) \in \mathbb{F}_2[X].$$

Plus généralement, si $n = \ell^r$ où ℓ est un nombre premier et $r \geqslant 1$, les

diviseurs de n sont de la forme ℓ^k, avec $k \in [\![0, r]\!]$. Or, on a

$$\mu(1) = 1, \quad \mu(\ell) = -1 \quad \text{et} \quad \mu(\ell^k) = 0, \quad \text{pour tout } k \in [\![2, r]\!].$$

Par conséquent,

$$I(\ell^r, p) = \frac{p^{\ell^r} - p^{\ell^{r-1}}}{\ell^r}, \quad \text{pour tout nombre premier } \ell \text{ et tout } r \geqslant 1.$$

On a donc $I(4, 2) = 3$. Un polynôme unitaire f de degré 4 est soit irréductible, soit possède une racine, soit admet un facteur irréductible de degré 2. Autrement dit, f sera irréductible si, et seulement si, f n'a pas de racines dans $\mathbb{F}_2$, et $X^2 + X + \overline{1} \nmid f$. On trouve alors les polynômes

$$X^4 + X^3 + \overline{1}, \quad X^4 + X + \overline{1}, \quad X^4 + X^3 + X^2 + X + \overline{1} \quad \in \mathbb{F}_2[X].$$

Avant de passer au premier résultat important de ce paragraphe, nous avons besoin d'un dernier lemme.

6.9. Lemme. *Tout morphisme d'anneaux entre deux corps finis de même cardinal est un isomorphisme*[5]. *Si de plus p est la caractéristique de ces deux corps finis, cet isomorphisme est un isomorphisme de $\mathbb{F}_p$-algèbres.*

Démonstration. Soient K et K' deux corps finis à p^n éléments. On a donc

$$p = \mathrm{car}(K) = \mathrm{car}(K').$$

Soit $\varphi : K \longrightarrow K'$ un morphisme d'anneaux. C'est en particulier un morphisme de groupes additifs. De plus, pour tout $m \in [\![0, p-1]\!]$, on a

$$\varphi(m{\cdot}x) = \varphi(x + \cdots + x) = \varphi(x) + \cdots + \varphi(x) = m{\cdot}\varphi(x).$$

Par définition de la loi externe sur K et K', cela se récrit

$$\varphi(\overline{m}{\cdot}x) = \overline{m}{\cdot}\varphi(x), \quad \text{pour tout } \overline{m} \in \mathbb{F}_p.$$

Ainsi, φ est $\mathbb{F}_p$-linéaire. Enfin, le noyau de φ est un idéal de K, donc est égal à (0) ou K. Dans le second cas, on aurait

$$1_{K'} = \varphi(1_K) = 0_{K'},$$

ce qui est impossible puisqu'un corps est non trivial. Par conséquent, φ est injective, et comme K et K' ont même nombre d'éléments, φ est bijective. Cela achève la démonstration. □

On en vient maintenant au théorème d'existence et d'unicité des corps finis.

5. Attention ! Nous ne prétendons pas pour l'instant qu'un tel morphisme existe nécessairement.

6.10. Théorème (Structure des corps finis). *Soit $q = p^n$, où p est un nombre premier et $n \geqslant 1$. Alors, il existe un corps fini à q éléments, unique à isomorphisme près. De plus, tout morphisme entre deux corps finis à q éléments est un isomorphisme de $\mathbb{F}_p$-algèbres.*

Démonstration. D'après le dernier point du corollaire 6.7, il existe un polynôme irréductible unitaire $f \in \mathbb{F}_p[X]$ de degré n. Soit $K = \mathbb{F}_p[X]/(f)$. D'après le lemme 6.2, c'est un corps fini à q éléments. Soit K' un autre corps fini à q éléments. D'après le lemme 6.3, on a $x^q - x = 0$, pour tout $x \in K'$. Comme $|K'| = q = \deg(X^q - X)$, on a

$$X^q - X = \prod_{x \in K'} (X - x) \in K'[X].$$

Le morphisme d'anneaux

$$\begin{aligned} \mathbb{F}_p &\longrightarrow K' \\ \overline{m} &\longmapsto m \cdot 1_{K'} \end{aligned}$$

induit un morphisme d'anneaux $\iota : \mathbb{F}_p[X] \longrightarrow K'[X]$.

Comme $f \mid X^q - X \in \mathbb{F}_p[X]$, $\iota(f) \in K'[X]$ divise $\iota(X^q - X) = X^q - X$ dans l'anneau $K'[X]$. D'après ce qui précède, $\iota(f)$ est scindé dans K', et possède donc une racine $\alpha \in K'$. Le morphisme $\varphi : \mathbb{F}_p[X] \longrightarrow K'$ obtenu par composition du morphisme ι et du morphisme d'évaluation

$$\begin{aligned} K'[X] &\longrightarrow K' \\ P &\longmapsto P(\alpha) \end{aligned}$$

vérifie donc $f \in \mathrm{Ker}(\varphi)$, et induit alors un morphisme $K \longrightarrow K'$. Ce morphisme est un isomorphisme de $\mathbb{F}_p$-algèbres par le lemme 6.9. Par conséquent, tout corps fini est isomorphe à K, ce qui achève la démonstration.□

Notation. Si q est une puissance non triviale d'un nombre premier, on note $\mathbb{F}_q$ l'unique corps à q éléments.

6.11. Remarque. Comme on l'a déjà remarqué auparavant, si $f \in \mathbb{F}_p[X]$ est un polynôme unitaire de degré $n \geqslant 1$, alors $\mathbb{F}_p[X]/(f)$ est un corps à p^n éléments. En particulier, $\mathbb{F}_q \simeq \mathbb{F}_p[X]/(f)$, et le choix d'un autre polynôme f donne le même corps à isomorphisme près.

6.12. Exemple. Puisque $X^3 + X + \overline{1} \in \mathbb{F}_2[X]$ et $X^3 + X^2 + \overline{1} \in \mathbb{F}_2[X]$ sont irréductibles, on a

$$\mathbb{F}_8 \simeq \mathbb{F}_2[X]/(X^3 + X + \overline{1}) \simeq \mathbb{F}_2[X]/(X^3 + X^2 + \overline{1}).$$

Il est très instructif d'essayer de construire un isomorphisme explicite entre les deux anneaux quotients. D'après le lemme 6.9, il suffit de trouver un morphisme $\varphi : \mathbb{F}_2[X]/(X^3 + X + \overline{1}) \longrightarrow \mathbb{F}_2[X]/(X^3 + X^2 + \overline{1})$.

D'après le théorème de factorisation, un tel morphisme est induit par un morphisme $f : \mathbb{F}_2[X] \longrightarrow \mathbb{F}_2[X]/(X^3+X^2+\overline{1})$ tel que $X^3+X+\overline{1} \in \mathrm{Ker}(f)$.

Un tel morphisme d'anneaux est clairement $\mathbb{F}_2$-linéaire. Il reste donc à choisir $\alpha = f(X)$ convenablement. Mais $f(X^3+X+\overline{1})$ est nul si, et seulement si, on a

$$\alpha^3 + \alpha + \overline{1} = \overline{0} \in \mathbb{F}_2[X]/(X^3+X^2+\overline{1}).$$

On doit donc trouver une racine du polynôme $T^3+T+\overline{1}$ dans l'anneau quotient $\mathbb{F}_2[X]/(X^3+X^2+\overline{1})$.

Soit $\beta \in \mathbb{F}_2[X]/(X^3+X^2+\overline{1})$ la classe de X. On sait que la famille $(\overline{1}, \beta, \beta^2)$ est une $\mathbb{F}_2$-base du quotient, et que

$$\beta^3 + \beta^2 + \overline{1} = \overline{0},$$

soit

$$\beta^3 = \beta^2 + \overline{1}.$$

Posons $\alpha = \beta + \overline{1}$. Alors,

$$\alpha^3 = \beta^3 + \beta^2 + \beta + \overline{1} = \beta,$$

et donc

$$\alpha^3 + \alpha + \overline{1} = \beta + (\beta + \overline{1}) + \overline{1} = \overline{0}.$$

Un isomorphisme entre les deux anneaux quotients est donc donné par le morphisme induit par l'évaluation

$$ev_{\beta+\overline{1}} : \mathbb{F}_2[X] \longrightarrow \mathbb{F}_2[X]/(X^3+X^2+\overline{1}).$$

Un problème crucial est donc de déterminer si un polynôme à coefficients dans $\mathbb{F}_p$ (ou même dans $\mathbb{F}_q$) est irréductible, d'une part pour construire les corps finis, et d'autre part pour étudier l'irréductibilité des polynômes à coefficients dans un anneau (via le critère de réduction modulo un idéal).

Nous allons maintenant donner des conditions nécessaires et suffisantes d'irréductibilité d'un polynôme $f \in \mathbb{F}_p[X]$.

6.13. Proposition. *Soit $f \in \mathbb{F}_p[X]$ un polynôme de degré $n \geqslant 1$. Alors, f est irréductible si, et seulement si, les conditions suivantes sont vérifiées :*

(1) *on a $f \mid X^{p^n} - X$;*

(2) *pour tout diviseur strict d de n, $\mathrm{pgcd}(f, X^{p^d} - X) = 1$.*

Démonstration. Soit $f \in \mathbb{F}_p[X]$. On peut toujours supposer f unitaire. Supposons que f soit irréductible. Alors, f vérifie (1) et (2) d'après le corollaire 6.6. Inversement, supposons que f vérifie les deux conditions de la proposition. Soit $\pi \in \mathbb{F}_p[X]$ un diviseur irréductible unitaire de f, et soit $d = \deg(\pi)$. La condition (1) montre alors que $d \mid n$. Supposons que d soit un diviseur strict de n. Alors, comme $\pi \mid X^{p^d} - X$, cela contredit (2). On a donc $d = n$, et ainsi $f = \pi$. Autrement dit, f est irréductible. □

Comme nous l'avons vu, il n'est pas vrai qu'un polynôme à coefficients dans un corps K sans racines dans K est irréductible. En revanche, on a le résultat suivant.

6.14. Proposition. *Soit $f \in \mathbb{F}_p[X]$ un polynôme de degré $n \geqslant 1$. Alors, f est irréductible si, et seulement si, pour tout $m \leqslant \frac{n}{2}$, f n'a pas de racines dans un corps fini à p^m éléments*[6].

Démonstration. Supposons que f soit réductible. Il possède donc un diviseur irréductible unitaire $\pi \in \mathbb{F}_p[X]$ de degré $m \leqslant \frac{n}{2}$. Soit $K = \mathbb{F}_p[X]/(\pi)$. Alors, K est un corps fini à p^m éléments, et π possède au moins une racine dans K, à savoir $\overline{X}$. Il en est donc de même de f.

Supposons que f soit irréductible, et qu'il existe un corps K à p^m éléments dans lequel f a une racine α, avec $m \leqslant \frac{n}{2}$. Alors, $K' = \mathbb{F}_p[X]/(f)$ est un corps à p^n éléments, et le morphisme $\mathbb{F}_p[X] \longrightarrow K$ d'évaluation en α induit un morphisme $K' \longrightarrow K$, nécessairement injectif, car K' et K sont des corps. En particulier, on a $p^n \leqslant p^m$, d'où une contradiction puisque l'on a $m \leqslant \frac{n}{2} < n$. Cela achève la démonstration. □

6.15. Remarque. Soient K et K' deux corps à p^m éléments. Comme deux tels corps sont isomorphes et qu'un isomorphisme est $\mathbb{F}_p$-linéaire, on en déduit aisément qu'un polynôme $f \in \mathbb{F}_p[X]$ a une racine dans K si, et seulement si, f a une racine dans K'.

Pour appliquer le critère précédent, il suffit de donc choisir sa construction explicite préférée de $\mathbb{F}_{p^m}$ et de vérifier l'existence ou non de racines dans cet unique cas.

6.16. Exemple. Soit $f = X^5 - X^2 - \overline{2} \in \mathbb{F}_5[X]$. Montrons qu'il est irréductible. On doit vérifier que f n'a pas de racines dans $\mathbb{F}_5$ et $\mathbb{F}_{25}$. Comme $\mathbb{F}_5$ s'identifie à un sous-anneau de $\mathbb{F}_{25}$, il suffit de vérifier que f n'a pas de racines dans $\mathbb{F}_{25}$. Le polynôme $X^2 - \overline{2} \in \mathbb{F}_5[X]$ n'a pas de racines dans $\mathbb{F}_5$, et donc $K = \mathbb{F}_5[X]/(X^2 - \overline{2})$ est isomorphe à $\mathbb{F}_{25}$.

Autrement dit, $\mathbb{F}_{25}$ est un $\mathbb{F}_5$-espace vectoriel de dimension 2, de base $(1, \alpha)$, où $\alpha \in \mathbb{F}_{25}$ vérifie $\alpha^2 = \overline{2}$.

Tout élément de $\mathbb{F}_{25}$ est donc de la forme $a + b\alpha$, $a, b \in \mathbb{F}_5$. En utilisant le lemme XV-1.18 et la relation $\alpha^2 = 2$, on obtient alors

$$(a + b\alpha)^5 = a^5 + b^5\alpha^5 = a + b\alpha^5 = a + 4b\alpha.$$

On a aussi

$$(a + b\alpha)^2 = a^2 + 2ab\alpha + b^2\alpha^2 = a^2 + 2b^2 + 2ab\alpha.$$

6. Ici, on identifie de manière légèrement abusive $\mathbb{F}_p$ avec le sous-anneau engendré par $1_{\mathbb{F}_{p^m}}$.

Finalement, on obtient

$$f(a+b\alpha) = (a - a^2 - 2b^2 - 2) + 2b(2-a)\alpha.$$

Comme $(1, \alpha)$ est une base, pour que $f(a+b\alpha) = 0$, il faut que $2b(2-a) = 0$. On a donc $a = 2$ ou $b = 0$.

Si $b = 0$, on a $f(a+b\alpha) = a - a^2 - 2$. On vérifie, en étudiant les cinq valeurs possibles pour a, que cette quantité est toujours non nulle.

Si $a = 2$, on a $f(a + b\alpha) = (-2b^2 - 4) = -2(b^2 + 2)$. Encore une fois, on vérifie que cela est toujours non nul.

Ainsi, f n'a pas de racines dans $\mathbb{F}_{25}$ et est donc irréductible.

On va maintenant étudier un algorithme de factorisation des polynômes dans $\mathbb{F}_p[X]$.

Soit $f = \sum_m a_m X^m \in \mathbb{F}_p[X]$ unitaire.

Remarquons tout d'abord que $f = g^p \in \mathbb{F}_p[X]$ si, et seulement si, $a_j = 0$ dès que $p \nmid j$.

En effet, si $f = g^p$, avec $g = \sum_m b_m X^m$, alors

$$g^p = \sum_m b_m^p X^{mp}.$$

Inversement, si $f = \sum_m b_{mp} X^{mp}$, alors on a

$$f = \sum_m b_{mp} X^{mp} = \sum_m b_{mp}^p X^{mp} = \Big(\sum_m b_{mp} X^m \Big)^p.$$

Soit maintenant $f \in \mathbb{F}_p[X]$ non nul. En regroupant les facteurs irréductibles dont l'exposant dans f est un multiple de p, on peut écrire f sous la forme

$$f = f_1^{e_1} \cdots f_r^{e_r} g^p,$$

avec $r \geqslant 0$, $e_i \geqslant 1$, $p \nmid e_i$ et $f_1, \ldots, f_r$ irréductibles unitaires tels que $f_i \nmid g$.

Alors, on vérifie facilement que $\operatorname{pgcd}(f, f') = f_1^{e_1 - 1} \cdots f_r^{e_r - 1} g^p$ (on prend le pgcd normalisé). Ainsi,

$$\frac{f}{\operatorname{pgcd}(f, f')} = f_1 \cdots f_r.$$

Un polynôme de la forme précédente $f_1 \cdots f_r$ où les f_i sont irréductibles et distincts deux à deux sera dit sans facteurs carrés. Remarquons que f est sans facteurs carrés si, et seulement si, $\operatorname{pgcd}(f, f') = 1$.

Supposons donc que $f = f_1 \cdots f_r$. Alors, le théorème chinois donne un isomorphisme de $\mathbb{F}_p$-algèbres

$$\varphi : \mathbb{F}_p[X]/(f) \xrightarrow{\sim} \mathbb{F}_p[X]/(f_1) \times \cdots \times \mathbb{F}_p[X]/(f_r).$$

Comme $P(X^p) = P(X)^p$, pour tout $P \in \mathbb{F}_p[X]$, le morphisme de $\mathbb{F}_p$-algèbres

$$\begin{aligned} u : \mathbb{F}_p[X]/(f) &\longrightarrow \mathbb{F}_p[X]/(f) \\ \overline{P} &\longmapsto \overline{P(X^p)} \end{aligned}$$

est alors bien défini, puisque $u(f) = f(X^p) = f^p \in (f)$.

6.17. Lemme. *Soit $P \in \mathbb{F}_p[X]$. Alors, on a $\overline{P} \in \mathrm{Ker}(u-\mathrm{Id})$ si, et seulement si, $\varphi(\overline{P}) = (\overline{\lambda}_1, \ldots, \overline{\lambda}_r)$ pour certains $\lambda_1, \ldots, \lambda_r \in \mathbb{F}_p$.*

En particulier, $r = \dim_{\mathbb{F}_p}\big(\mathrm{Ker}(u - \mathrm{Id})\big) = \deg(f) - \dim_{\mathbb{F}_p}\big(\mathrm{Im}(u - \mathrm{Id})\big)$.

Démonstration. Puisque $Y^p - Y = \prod_{\lambda \in \mathbb{F}_p} (Y - \lambda)$, en évaluant Y en $P(X)$, on obtient

$$P(X^p) - P(X) = P(X)^p - P(X) = \prod_{\lambda \in \mathbb{F}_p} \big(P(X) - \lambda\big).$$

On a donc $\overline{P} \in \mathrm{Ker}(u - \mathrm{Id}) \iff \prod_{\lambda \in \mathbb{F}_p} \big(P(X) - \lambda\big) \equiv 0\ [f]$.

Si $\overline{P} \in \mathrm{Ker}(u - \mathrm{Id})$, alors pour tout i, on obtient en particulier

$$\prod_{\lambda \in \mathbb{F}_p} \big(P(X) - \lambda\big) \equiv 0\ [f_i].$$

Comme $\mathbb{F}_p[X]/(f_i)$ est un corps, puisque f_i est irréductible, par intégrité, il existe $\lambda_i \in \mathbb{F}_p$ tel que $P(X) \equiv \lambda_i\ [f_i]$.

Inversement, si $P(X) \equiv \lambda_i\ [f_i]$ pour tout $i \in [\![1, r]\!]$, avec $\lambda_i \in \mathbb{F}_p$, alors

$$\prod_{\lambda \in \mathbb{F}_p} \big(P(X) - \lambda\big) \equiv 0\ [f_i], \quad \text{pour tout } i \in [\![1, r]\!].$$

Comme les f_i sont irréductibles et distincts deux à deux, on en déduit que

$$\prod_{\lambda \in \mathbb{F}_p} \big(P(X) - \lambda\big) \equiv 0\ [f],$$

c'est-à-dire $P(X)^p \equiv P(X)\ [f]$, ou encore $\overline{P} \in \mathrm{Ker}(u - \mathrm{Id})$. Le dernier point est alors clair. □

Remarquons que l'on peut alors calculer le nombre de facteurs irréductibles de f grâce à l'algèbre linéaire. On écrit la matrice représentative M de u dans la base $\big(\overline{1}, \overline{X}, \ldots, \overline{X^{\deg(f)-1}}\big)$ en réduisant les monômes

$$1,\ X^p,\ X^{2p},\ \ldots,\ X^{(\deg(f)-1)p}$$

modulo f. Le nombre de facteurs irréductibles de f dans $\mathbb{F}_p[X]$ est alors égal à $\dim_{\mathbb{F}_p}\big(\mathrm{Ker}(M - \mathrm{I}_n)\big)$, où $n = \deg(f)$.

Avant de pouvoir décrire l'algorithme, on a besoin d'un dernier lemme.

6.18. Lemme. *Soit $P \in \mathbb{F}_p[X]$ tel que $\overline{P} \in \mathrm{Ker}(u - \mathrm{Id})$. Alors,*

$$f = \prod_{\lambda \in \mathbb{F}_p} \mathrm{pgcd}\big(f, (P(X) - \lambda)\big).$$

Démonstration. Appelons g le membre de droite. Remarquons que l'ensemble des diviseurs irréductibles unitaires de g est inclus dans l'ensemble des diviseurs irréductibles unitaires de f. Puisque $\overline{P} \in \mathrm{Ker}(u - \mathrm{Id})$, d'après le lemme précédent, on a $P \equiv \lambda_i \mod (f_i)$ pour tout entier $i \in [\![1, r]\!]$, avec $\lambda_i \in \mathbb{F}_p$. Fixons i. Si $\lambda \neq \lambda_i$, alors $P - \lambda$ n'est pas divisible par f_i. Sinon, on aurait

$$\lambda \equiv P \equiv \lambda_i \mod (f_i),$$

et donc on aurait $\lambda = \lambda_i$, d'où une contradiction. Si $\lambda = \lambda_i$, alors $f_i \mid P - \lambda$ par définition de λ_i, et comme il divise f, alors $f_i \mid \mathrm{pgcd}(f, P - \lambda_i)$ et de plus $f_i^2 \nmid \mathrm{pgcd}(f, P - \lambda_i)$ (car $f_i^2 \nmid f$). On en déduit que $f_i \mid g$ et $f_i^2 \nmid g$. Comme f et g sont unitaires sans facteurs carrés et ont mêmes diviseurs irréductibles, on en déduit que $f = g$. □

On peut maintenant décrire l'algorithme de Berlekamp.

Algorithme de Berlekamp

Étape 1. On calcule f'.

Si $f' = 0$, on écrit $f = g^p$ avec $g \in \mathbb{F}_p[X]$, et l'on applique l'algorithme à g en repartant du début.

Étape 2. Si $f' \neq 0$, on calcule $\mathrm{pgcd}(f, f')$. Si $\mathrm{pgcd}(f, f') \neq 1$, on applique l'algorithme du début à chacun des polynômes $\mathrm{pgcd}(f, f')$ et $\dfrac{f}{\mathrm{pgcd}(f, f')}$.

Étape 3. Si $\mathrm{pgcd}(f, f') = 1$, f est donc un polynôme sans facteurs carrés. On calcule alors r comme expliqué précédemment. Si $r = 1$, f est irréductible.

Étape 4. Si $r \geqslant 2$, on choisit un élément $\overline{P} \in \mathrm{Ker}(u - \mathrm{Id})$, avec P non constant et de degré $< \deg(f)$ (on peut en calculer un explicitement grâce à la matrice M). Alors, on a $1 \leqslant \deg(P - \lambda) < \deg(f)$, pour tout $\lambda \in \mathbb{F}_p$.

Étape 5. On utilise le lemme précédent pour factoriser f. La factorisation obtenue est non triviale. Les facteurs obtenus sont encore sans facteurs carrés, et l'on repart à l'étape 1 pour chacun des facteurs.

Au bout d'un nombre fini d'étapes, on aura factorisé f.

On a en particulier une méthode pour décider si un polynôme $f \in \mathbb{F}_p[X]$ de degré $n \geqslant 1$ est irréductible ou non.

Algorithme

(1) Si $\operatorname{pgcd}(f, f') \neq 1$, alors f n'est pas irréductible [7].

(2) Si $\operatorname{pgcd}(f, f') = 1$, alors f est sans facteurs carrés. Pour $j \in [\![0, n-1]\!]$, on calcule le reste R_j de la division de X^{jp} par f, que l'on écrit

$$R_j = \sum_{i=0}^{n-1} a_{ij} X^i, \quad a_{ij} \in \mathbb{F}_p.$$

(3) On pose $M = (a_{ij}) \in \mathrm{M}_n(\mathbb{F}_p)$, et on calcule $r = \dim_{\mathbb{F}_p} \big(\operatorname{Ker}(M - \mathrm{I}_n)\big)$.

(4) Le polynôme f est irréductible si, et seulement si, $r = 1$.

6.19. Remarques

(1) Il est assez remarquable d'avoir un algorithme efficace pour factoriser systématiquement un polynôme à coefficients dans $\mathbb{F}_p$, ou décider de l'irréductibilité d'un polynôme. Notons qu'un tel algorithme n'est pas disponible à l'heure actuelle pour les entiers.

(2) Tous les résultats de ce paragraphe sont vrais si l'on remplace $\mathbb{F}_p$ par un corps fini $\mathbb{F}_q$. Par exemple, il existe toujours un polynôme unitaire irréductible de $\mathbb{F}_q[X]$ de degré fixé, et tous les critères d'irréductibilité, ainsi que l'algorithme de Berlekamp sont encore valables (cf. [7], par exemple).

7. Exercices

E1. a. Soit $P \in \mathbb{Z}[X]$ un polynôme tel que $|P(0)|$ est un nombre premier. Si P n'a pas de racines complexes de module $\leqslant 1$, montrer que P est irréductible dans $\mathbb{Z}[X]$.

Indication. Si P est réductible, montrer que P possède un facteur Q non constant tel que $Q(0) = 1$. Considérer alors le produit des racines de Q, et en déduire une contradiction.

b. Soit p un nombre premier et n un entier $\geqslant 1$. Montrer que le polynôme $P = X^n + X^{n-1} + \cdots + X + p$ est irréductible dans $\mathbb{Z}[X]$ et $\mathbb{Q}[X]$.

Indication. Soit $z \in \mathbb{C}$ une racine de P de module $\leqslant 1$. Calculer z^{n+1} en fonction de z et p, et montrer que $\operatorname{Re}(z) \geqslant 1$, avec égalité si, et seulement si, $z = 1$.

E2. Soient $a_1, \ldots, a_n \in \mathbb{Z}$ deux à deux distincts ($n \geqslant 1$).

7. La vérification est laissée au lecteur.

Montrer que le polynôme $f = (X - a_1) \cdots (X - a_n) - 1$ est irréductible dans $\mathbb{Z}[X]$ et $\mathbb{Q}[X]$.

Indication. Si $f = gh \in \mathbb{Z}[X]$, montrer que l'on peut supposer g et h unitaires. Vérifier que pour tout $i \in [\![1, n]\!]$, il existe $\varepsilon_i \in \{\pm 1\}$ tel que

$$g(a_i) = \varepsilon_i \quad \text{et} \quad h(a_i) = -\varepsilon.$$

Montrer alors que $g + h = 0$, et conclure.

E3. Démontrer les formules énoncées dans la remarque 5.8.

E4. Le but de cet exercice est de déterminer quels sont les polynômes unitaires de $\mathbb{Z}[X]$ dont toutes les racines complexes sont dans le disque-unité fermé.

On suppose tout d'abord que $P \in \mathbb{Z}[X]$ est un polynôme irréductible unitaire, distinct de X, dont toutes les racines sont dans le disque-unité fermé. On veut montrer que P est un polynôme cyclotomique.

a. En considérant $|P(0)|$, montrer que toutes les racines de P sont de module 1.

b. Soit $n = \deg(P)$, et soient $z_1, \ldots, z_n \in \mathbb{C}$ les racines de P (non nécessairement distinctes). Pour $m \geqslant 1$, on pose

$$P_m = \prod_{i=1}^{n} (X - z_i^m).$$

En utilisant le théorème XV-7.11, montrer que $P_m \in \mathbb{Z}[X]$.

c. Montrer que les coefficients de P_m sont bornés en valeur absolue par une constante ne dépendant que de n.

d. En déduire que les ensembles $R_i = \{z_i^m \mid m \geqslant 1\}$ sont finis.

e. Conclure.

On revient au cas général.

f. Montrer le théorème de Kronecker : un polynôme $P \in \mathbb{Z}[X]$ unitaire a toutes ses racines dans le disque-unité si, et seulement si, on a

$$P = X^m \Phi_{n_1}^{r_1} \cdots \Phi_{n_s}^{r_s},$$

avec m, $r_i \geqslant 0$ et $n_i \geqslant 1$.

E5. Justifier que l'on a un isomorphisme d'anneaux

$$\mathbb{F}_5[X]/(X^2 - 2) \simeq \mathbb{F}_5[X]/(X^2 + X + 1),$$

et trouver un isomorphisme explicite.

E6. Déterminer parmi les anneaux suivants lesquels sont isomorphes :

$$\mathbb{F}_3[X]/(X^2+1), \quad \mathbb{F}_3[X]/(X^2+X+2), \quad \mathbb{Z}/9\mathbb{Z}, \quad \mathbb{Z}/3\mathbb{Z}\times\mathbb{Z}/3\mathbb{Z}.$$

On justifiera avec soin (mais avec un minimum de calculs) pourquoi deux anneaux dans la liste sont isomorphes ou non. Pour chaque paire d'anneaux isomorphes dans la liste ci-dessus, on explicitera un isomorphisme.

E7. Étudier l'irréductibilité du polynôme $P \in A[X]$, dans chacun des cas suivants :

(i) $A=\mathbb{Z}$, $P=X^3-14789X+245798421$;
(ii) $A=\mathbb{R}$, $P=X^5+2X^2+2$;
(iii) $A=\mathbb{F}_5$, $P=X^5+2X^2+2$;
(iv) $A=\mathbb{Q}$, $P=X^n+m(m+1)X^r+2$, $r\in[\![1,n-1]\!]$, $m\in\mathbb{Z}$;
(v) $A=\mathbb{Q}(i)$, $P=X^{18}-2X+(1-i)$;
(vi) $A=\mathbb{F}_2$, $P=X^5-X^2+\overline{1}$;
(vii) $A=\mathbb{F}_2$, $P=X^7+X^4-\overline{1}$;
(viii) $A=\mathbb{Q}$, $P=X^5-35X^2+11$;
(ix) $A=\mathbb{Q}$, $f=X^{p^m}+p-1$, p premier, $m\geqslant 1$.

E8. Soit A un anneau commutatif **intègre**.

a. Pour $P\in A[X]$, on pose

$$P^*=X^{v_X(P)+\deg(P)}P\Big(\frac{1}{X}\Big),$$

avec la convention $0^*=0$. Vérifier que $P^*\in A[X]$.

b. Montrer que P est irréductible dans $A[X]$ si, et seulement si, P^* est irréductible dans $A[X]$.

c. Soit p un nombre premier. Étudier l'irréductibilité dans $\mathbb{Q}[X]$ du polynôme

$$P=\sum_{k=0}^{p}\frac{X^k}{(p-k)!}.$$

E9. Soit $q=p^n$, p premier, $n\geqslant 1$. Calculer

$$S=\sum_{x\in\mathbb{F}_q}x \quad \text{et} \quad P=\prod_{x\in\mathbb{F}_q^\times}x.$$

On distinguera les cas $q=2$ et $q>2$.

E10. Soit p un nombre premier, et soient d, $n \geqslant 1$ deux entiers. Montrer que $\mathbb{F}_{p^d}$ est isomorphe à un sous-corps de $\mathbb{F}_{p^n}$ si, et seulement si, $d \mid n$.

Indication. Pour « $\Longrightarrow$ », on pourra utiliser le fait que si K est un corps fini, $K^\times$ est un groupe cyclique. Si $d \mid n$, construire un morphisme d'anneaux $\mathbb{F}_p[X]/(f) \longrightarrow \mathbb{F}_{p^n}$ où $f \in \mathbb{F}_p[X]$ est un polynôme irréductible unitaire de degré d.

E11. Soit $\Phi_8 = X^4 + 1 \in \mathbb{Z}[X]$.

a. Calculer $\Phi_8(X+1)$ et retrouver le fait que Φ_8 est irréductible dans l'anneau $\mathbb{Z}[X]$.

b. Montrer que Φ_8 est réductible modulo p pour tout p premier.

c. Montrer que Φ_8 est irréductible modulo 4.

Indication. Par l'absurde, supposer que Φ_8 est réductible modulo 4. Utiliser le corollaire 4.13 et la réduction modulo 2 pour obtenir une contradiction.

E12. On suppose connus les définitions et résultats de l'exercice XVIII-20. Montrer qu'un anneau à pgcd est intégralement clos.

E13. On suppose connus les définitions et résultats de l'exercice XVIII-20.

a. Soit $A = \mathbb{Z}[i\sqrt{6}]$. Montrer que $f = 2X + i\sqrt{6}$ et $g = 3X + i\sqrt{6}$ sont deux polynômes primitifs de $A[X]$, mais que fg n'est pas primitif[8].

On se propose maintenant de montrer que si A est un anneau à pgcd, le produit de deux polynômes primitifs est primitif.

b. Soit $h \in A[X]$ et soit $a \in A$. Montrer que si d est un pgcd des coefficients de h, alors ad est un pgcd des coefficients de ah.

c. Soient f, $g \in A[X]$ deux polynômes primitifs. On suppose que fg n'est pas primitif. Soit $a \in A$ un pgcd des coefficients de fg.

(i) Justifier que a est non nul et non inversible, et en déduire que le polynôme $\overline{f} \in A/(a)[X]$ est un diviseur de zéro.

(ii) En déduire qu'il existe $d \in A$ et $h \in A[X]$ tels que $a \nmid d$ et $df = ah$.

(iii) En utilisant **b.**, aboutir à une contradiction, et conclure.

E14. Soit A un anneau commutatif, et soit $a \in A \setminus \{0\}$.

8. Merci à Vincent Beck pour ce contre-exemple.

a. On suppose que A est intègre. Montrer que X et $X-a$ sont premiers entre eux, et que pour tout $P \in A[X]$, on a

$$X \mid P \text{ et } X-a \mid P \Longrightarrow X(X-a) \mid P.$$

b. On suppose dans cette question que $A = \mathbb{Z}/4\mathbb{Z}$.

(i) Montrer que les polynômes X et $X-\overline{2}$ sont irréductibles et non associés dans $A[X]$.

(ii) En déduire que X et $X-\overline{2}$ sont premiers entre eux dans $A[X]$.

(iii) Montrer que $\overline{2}X$ est divisible par X et par $X-\overline{2}$ dans $A[X]$, mais qu'il n'est pas divisible par $X(X-\overline{2})$.

Chapitre XXII

Idéaux de $A[X]$

Dans ce chapitre, on s'intéresse de plus près aux idéaux de $A[X]$. En particulier, on cherche à savoir sous quelles conditions ces idéaux sont de type fini, voire principaux. Tous les anneaux considérés seront supposés commutatifs.

1. Principalité des idéaux de $A[X]$

Lorsque K est un corps, l'anneau $K[X]$ est euclidien, donc principal. La réciproque est en fait vraie, comme le montre le résultat suivant.

1.1. Proposition. *Soit A un anneau. Les propriétés suivantes sont équivalentes :*

(1) *l'anneau A est un corps ;*

(2) *l'anneau $A[X]$ est euclidien ;*

(3) *l'anneau $A[X]$ est principal.*

Démonstration. L'implication (1) $\Longrightarrow$ (2) a déjà été constatée, et l'implication (2) $\Longrightarrow$ (3) vient de la proposition XVIII-3.3. Supposons maintenant que $A[X]$ soit principal. En particulier, $A[X]$ est intègre, et donc A est également intègre par le lemme XV-6.9. Or, l'évaluation en 0 est un morphisme d'anneaux surjectif, de noyau (X), et par le théorème de factorisation, on obtient un isomorphisme d'anneaux

$$A[X]/(X) \simeq A.$$

Par conséquent, l'anneau $A[X]/(X)$ est intègre et (X) est donc un idéal premier non nul (car X est non nul) d'après la proposition XVIII-2.7. L'anneau $A[X]$ étant supposé principal, (X) est alors un idéal maximal par

la proposition XVIII-2.7. Ainsi, $A[X]/(X)$ est un corps, et par suite, A est un corps. □

Lorsque A est intègre, on peut démontrer cette proposition de manière directe en exhibant un idéal de $A[X]$ non principal lorsque A n'est pas un corps.

1.2. Exemple. Soit A un anneau intègre qui n'est pas un corps. Alors, pour tout élément $a \in A$ non nul et non inversible, l'idéal $\mathfrak{a} = (a, X)$ n'est pas principal.

Remarquons tout d'abord que l'on a l'égalité

$$\mathfrak{a} = \{f \in A[X] \mid f(0) \in (a)\}.$$

En effet, l'inclusion « $\subset$ » est claire, et l'autre inclusion provient du fait que tout polynôme f s'écrit

$$f = f(0) + Xg, \quad g \in A[X].$$

Supposons maintenant que l'on ait $(a, X) = (f)$ pour un certain $f \in \mathfrak{a}$. Puisque a est un élément de $\mathfrak{a}$, on a $a = fg$, $g \in A[X]$. Puisque A est intègre, on en déduit que $\deg(f) + \deg(g) = 0$.

Ainsi, $\deg(f) = 0$, et donc $f = b$, $b \in A$. Comme $X \in (a, X) = (b)$, on en déduit que $X = bh$, $h \in A[X]$. En comparant les degrés, on voit que h est de degré 1, puis en comparant les termes en X, on en déduit que $b \in A^\times$. Il s'ensuit que $(a, X) = (b) = A[X]$. Il existe donc $u, v \in A[X]$ tels que

$$au + Xv = 1.$$

En évaluant en 0, on obtient $a \in A^\times$, d'où une contradiction.

Finalement, (a, X) est un idéal non principal.

L'hypothèse d'intégrité de A est essentielle pour obtenir concrètement un idéal non principal de $A[X]$ lorsque A n'est pas un corps. En effet, lorsque l'anneau A n'est pas intègre, il se peut très bien que tous les idéaux de l'anneau $A[X]$ soient principaux malgré tout, comme nous allons le voir maintenant (ce qui ne contredit nullement la proposition 1.1, puisqu'un anneau principal est intègre).

Le but de ce qui suit est de démontrer le résultat suivant.

1.3. Théorème. *Soit A un anneau. Alors, tout idéal de $A[X]$ est principal si, et seulement si, A est isomorphe à un produit fini de corps.*

On s'occupe d'abord du sens facile.

1.4. Lemme. *Soit A un anneau. On suppose que A est isomorphe à un produit fini de corps. Alors, tout idéal de $A[X]$ est principal.*

Démonstration. Supposons que $A \simeq K_1 \times \cdots \times K_r$, où K_i est un corps. On a alors

$$A[X] \simeq (K_1 \times \cdots \times K_r)[X] \simeq K_1[X] \times \cdots \times K_r[X].$$

Ces isomorphismes sont faciles à établir, et laissés à la sagacité du lecteur. Puisque K_i est un corps, $K_i[X]$ est principal, et donc tous ses idéaux sont principaux. Par le lemme XVIII-4.4, tout idéal de $K_1[X] \times \cdots \times K_r[X]$ est alors principal. Il s'ensuit aisément que tout idéal de $A[X]$ est principal (exercice facile). □

Ainsi, on a établi le sens direct. L'autre sens est plus délicat, et nous aurons besoin de plusieurs résultats intermédiaires.

1.5. Lemme. *Soit A un anneau tel que tout idéal de $A[X]$ soit principal. Alors, tout idéal premier de A est maximal.*

Démonstration. Soit $\mathfrak{p}$ un idéal premier de A. Alors, $A/\mathfrak{p}$ est intègre, donc l'anneau $(A/\mathfrak{p})[X]$ aussi. Or, on a un isomorphisme d'anneaux

$$(A/\mathfrak{p})[X] \simeq A[X]/\mathfrak{p}A[X].$$

Comme les idéaux de $A[X]/\mathfrak{p}A[X]$ sont exactement de la forme $I/\mathfrak{p}A[X]$, où I est un idéal de $A[X]$ contenant $\mathfrak{p}A[X]$, ceux-ci sont principaux par hypothèse. On en déduit alors que tous les idéaux de $(A/\mathfrak{p})[X]$ sont principaux, et comme c'est un anneau intègre, c'est exactement dire que $(A/\mathfrak{p})[X]$ est un anneau principal.

Ainsi, $A/\mathfrak{p}$ est un corps par la proposition 1.1, et $\mathfrak{p}$ est maximal. □

1.6. Lemme. *Soit A un anneau tel que tout idéal de $A[X]$ soit principal. Alors, tout idéal de A est principal. En particulier, A est noethérien.*

Démonstration. Soit $\mathfrak{a}$ un idéal de A. Alors, l'idéal $\mathfrak{A}$ de $A[X]$ engendré par les éléments de $\mathfrak{a}$ est principal par hypothèse, engendré par P. Montrons que $\mathfrak{a}$ est engendré par $P(0)$.

Soit en effet $b \in \mathfrak{a} \subset \mathfrak{A}$ et $Q \in A[X]$ un polynôme tel que $QP = b$. Alors, on a $b = P(0)Q(0) \in P(0)A$, donc $\mathfrak{a} \subset P(0)A$. Inversement, par définition, P est de la forme $\sum_i a_i P_i$, $a_i \in \mathfrak{a}$, $P_i \in A[X]$. En prenant les termes constants, on voit que $P(0) \in \mathfrak{a}$, donc finalement $\mathfrak{a} = P(0)A$. □

1.7. Lemme. *Soit A un anneau tel que tout idéal de $A[X]$ soit principal. Alors, l'anneau A n'a pas d'éléments nilpotents non nuls.*

Démonstration. Il suffit de montrer que pour tout $a \in A$ tel que $a^2 = 0$, on a $a = 0$.

En effet, supposons l'avoir démontré. Si $a \in \mathrm{Nil}(A)$, alors on a $a^n = 0$ pour un certain $n \geqslant 1$. Si $n = 1$, alors $a = 0$, et c'est gagné. Si $n \geqslant 2$, alors

$$(a^{n-1})^2 = a^{2n-2} = a^{n-2}a^n = 0,$$

et donc $a^{n-1} = 0$. Par récurrence descendante, on montre que $a = 0$.

Soit donc $a \in A$ tel que $a^2 = 0$ et soit $\mathfrak{a} = (a, X)$. Par hypothèse, $\mathfrak{a}$ est principal. Soit P un générateur de $\mathfrak{a}$.

Alors, $P = aR + XS$ et $X = QP$. En multipliant la première égalité par Q, on obtient donc $X = aQR + XQS$, soit $(1 - QS)X = aQR$.

Tous les coefficients de $(1-QS)X$ étant des multiples de a, il est en de même pour $1 - QS$ (il suffit d'identifier les coefficients). Ainsi, on a $QS = 1 + aT$ et donc $(1 - aT)QS = 1 - a^2T^2 = 1$.

Posons $U = (1 - aT)S$. On a alors $QU = 1$, et donc

$$P = UQP = UX.$$

Soit W tel que $a = WP$. Alors, $a = WUX$, et par évaluation en 0, on obtient $a = 0$. □

Nous aurons aussi besoin d'un résultat général sur les anneaux noethériens.

1.8. Proposition. *Soit A un anneau noethérien non trivial. Alors, tout idéal contient un produit fini non vide d'idéaux premiers. En particulier, il existe un produit fini d'idéaux premiers qui soit nul.*

Démonstration. Soit $\mathscr{F}$ l'ensemble des idéaux de A ne contenant pas un produit fini non vide d'idéaux premiers. Remarquons que $A \notin \mathscr{F}$ car, par le théorème de Krull, A contient un idéal maximal, donc premier. De plus, $\mathscr{F}$ ne contient aucun idéal premier.

Par la proposition XVIII-4.6, $\mathscr{F}$ possède un élément $\mathfrak{b}$ maximal pour l'inclusion. Montrons que $\mathfrak{b}$ est un idéal premier, ce qui donnera une contradiction. Par le point précédent, $\mathfrak{b} \neq A$. Supposons maintenant qu'il existe des éléments b_1, $b_2 \in A \setminus \mathfrak{b}$ tels que $b_1b_2 \in \mathfrak{b}$. Remarquons que l'on a

$$\big((b_1) + \mathfrak{b}\big)\big((b_2) + \mathfrak{b}\big) \subset (b_1b_2) + b_1\mathfrak{b} + b_2\mathfrak{b} \subset \mathfrak{b}.$$

Or, les idéaux $(b_i) + \mathfrak{b}$ contiennent strictement $\mathfrak{b}$, donc ne sont pas dans $\mathscr{F}$ par maximalité de $\mathfrak{b}$ pour l'inclusion. En particulier, ils contiennent un produit fini non vide d'idéaux premiers, et par suite $\big((b_1) + \mathfrak{b}\big)\big((b_2) + \mathfrak{b}\big)$ aussi. Mais alors, il en est de même de $\mathfrak{b}$ au vu de l'inclusion précédente, d'où une contradiction.

Ainsi, $\mathfrak{b}$ est un idéal premier, d'où encore une contradiction. L'ensemble $\mathscr{F}$ est par conséquent vide, et tout idéal contient un produit fini non vide d'idéaux premiers.

En particulier, c'est le cas pour l'idéal nul, d'où la dernière partie de la proposition. □

Démonstration du théorème 1.3.

Supposons que tous les idéaux de $A[X]$ soient principaux. Alors, A est noethérien par le lemme 1.6. Par la proposition 1.8, il existe des idéaux premiers $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$ distincts deux à deux et des entiers $m_1, \ldots, m_r$ non nuls tels que

$$\mathfrak{p}_1^{m_1} \cdots \mathfrak{p}_r^{m_r} = 0.$$

Si l'on pose $m = \max(m_i)$, on a $(\mathfrak{p}_1 \cdots \mathfrak{p}_r)^m = 0$.

L'idéal $\mathfrak{p}_1 \cdots \mathfrak{p}_r$ est principal par le lemme 1.6, et un générateur arbitraire est donc nilpotent, donc nul par le lemme 1.7, d'où $\mathfrak{p}_1 \cdots \mathfrak{p}_r = 0$. Par le lemme 1.5, chaque $\mathfrak{p}_i$ est maximal. Les idéaux $\mathfrak{p}_1, \ldots, \mathfrak{p}_r$ étant distincts deux à deux, ils sont alors comaximaux deux à deux.

En effet, pour tous $i, j \in [\![1, r]\!]$, $i \neq j$, il existe $a_{ij} \in \mathfrak{p}_j \setminus \mathfrak{p}_i$. Par conséquent, on a $\mathfrak{p}_i \subsetneq \mathfrak{p}_i + a_{ij}A$, et par maximalité de $\mathfrak{p}_i$, on en déduit que $\mathfrak{p}_i + a_{ij}A = A$. Mais alors, puisque $\mathfrak{p}_i + a_{ij}A \subset \mathfrak{p}_i + \mathfrak{p}_j$, on obtient

$$\mathfrak{p}_i + \mathfrak{p}_j = A, \quad \text{pour tous } i, j \in [\![1, r]\!],\ i \neq j.$$

Par le théorème chinois, on a alors

$$A \simeq A/(0) = A/\mathfrak{p}_1 \cdots \mathfrak{p}_r \simeq A/\mathfrak{p}_1 \times \cdots \times A/\mathfrak{p}_r.$$

Chaque $\mathfrak{p}_i$ étant maximal, $A/\mathfrak{p}_i$ est un corps pour tout $i \in [\![1, r]\!]$. La réciproque étant donnée par le lemme 1.4, cela achève la démonstration du théorème. □

1.9. Remarque. Le résultat du théorème 1.3 est bien connu, mais, curieusement, il est difficile de trouver une référence proposant une démonstration directe. Les arguments présentés ici sont de Christophe Chalons.

2. Noethérianité de $A[X]$

Nous allons maintenant nous intéresser aux conditions sous lesquelles $A[X]$ est noethérien.

On commence par un lemme facile.

2.1. Lemme. *Soit A un anneau. Si $A[X]$ est noethérien, alors A est noethérien.*

Démonstration. Puisque $A[X]$ est noethérien, l'anneau $A[X]/(X)$ est également noethérien par le lemme XVIII-4.3. Or, $A[X]/(X) \simeq A$, et donc A est également noethérien. □

Cette condition nécessaire est également suffisante.

2.2. Théorème (Hilbert). *Soit A un anneau. Alors, $A[X]$ est noethérien si, et seulement si, A est noethérien.*

Démonstration. Le sens direct étant donné par le lemme précédent, on s'occupe de la réciproque. Nous allons proposer deux démonstrations : une démonstration par l'absurde, plutôt courte, mais qui a l'inconvénient de ne pas être explicite, et une démonstration un peu plus longue, mais qui à l'avantage de produire (au moins en théorie) une famille génératrice finie d'un idéal de $A[X]$ donné.

Première démonstration. Soit A un anneau noethérien. On suppose que l'anneau $A[X]$ n'est pas noethérien, et qu'il existe un idéal $\mathfrak{a}$ de $A[X]$ qui ne soit pas de type fini.

Nous allons d'abord démontrer l'existence d'une suite $(f_m)_{m\geqslant 0}$ d'éléments de $\mathfrak{a}$ tel que, pour tout entier $m \geqslant 0$, f_m est un élément de degré minimal de $\mathfrak{a} \setminus (f_0, \ldots, f_{m-1})$. Puisque $\mathfrak{a}$ n'est pas de type fini, il est non nul. On peut donc choisir un polynôme $f_0 \in \mathfrak{a} \setminus \{0\}$ de degré minimal. Supposons maintenant que les polynômes $f_0, \ldots, f_m$ soient construits, pour un certain entier $m \geqslant 0$. Comme $\mathfrak{a}$ n'est pas de type fini, l'ensemble $\mathfrak{a} \setminus (f_0, \ldots, f_m)$ est non vide, et l'on peut choisir un élément f_{m+1} de cet ensemble de degré minimal. On construit ici la suite désirée par récurrence.

Remarquons que la suite d'entiers $\big(\deg(f_m)\big)_{m\geqslant 0}$ est croissante. En effet, supposons au contraire qu'il existe $m \geqslant 0$ tel que $\deg(f_{m+1}) < \deg(f_m)$. Comme $\mathfrak{a} \setminus (f_0, \ldots, f_m) \subset \mathfrak{a} \setminus (f_0, \ldots, f_{m-1})$, f_{m+1} serait alors un élément de $\mathfrak{a} \setminus (f_0, \ldots, f_{m-1})$ dont le degré est strictement inférieur à celui de f_m, ce qui contredit le choix même de f_m.

Pour tout $m \geqslant 0$, soit $a_m \in A \setminus \{0\}$ le coefficient dominant de f_m. Comme A est noethérien, la chaîne croissante d'idéaux

$$(a_0) \subset (a_0, a_1) \subset \cdots \subset (a_0, \ldots, a_{m-1}) \subset (a_0, \ldots, a_m) \subset \ldots$$

est stationnaire. En particulier, il existe un entier $n \geqslant 1$ tel que l'on ait

$$(a_0, \ldots, a_{n-1}) = (a_0, \ldots, a_n).$$

On peut donc écrire

$$a_n = u_0 a_0 + \cdots + u_{n-1} a_{n-1}, \quad u_0, \ldots, u_{n-1} \in A.$$

Posons maintenant $g = \sum_{i=0}^{n-1} u_i X^{\deg(f_n)-\deg(f_i)} f_i$. Notons que c'est bien un élément de $A[X]$, puisque $\deg(f_n) \geqslant \deg(f_i)$ pour tout $i \in [\![0, n-1]\!]$ d'après le point précédent. Par définition, on a $\deg(g) \leqslant \deg(f_n)$.

De plus, le coefficient correspondant à $X^{\deg(f_n)}$ est par construction égal à

$$u_0 a_0 + \cdots + u_{n-1} a_{n-1} = a_n \in A \setminus \{0\}.$$

Ainsi, $\deg(g) = \deg(f_n)$, et le coefficient dominant de g est égal à celui de f_n. On a donc $\deg(f_n - g) < \deg(f_n)$. Mais, $g \in (f_0, \ldots, f_{n-1})$. Par conséquent, $f_n - g$ est un élément de $\mathfrak{a} \setminus (f_0, \ldots, f_{n-1})$ dont le degré est strictement inférieur à celui de f_n, d'où une contradiction avec le choix de f_n. Ainsi, tout idéal de $A[X]$ est de type fini. Cela achève la démonstration.

Seconde démonstration. Soit A un anneau noethérien. Si $f \in A[X] \setminus \{0\}$, on note $\delta(f) \in A$ son coefficient dominant. Soit $\mathfrak{a}$ un idéal de $A[X]$. Pour tout $n \geqslant 0$, on pose

$$\mathcal{D}_n = \{\delta(f) \mid f \in \mathfrak{a},\ \deg(f) = n\} \cup \{0\}.$$

Commençons par montrer que $\mathcal{D}_n$ est un idéal de A. Soient x, $y \in \mathcal{D}_n$, et soit $a \in A$.

Par définition, $0 \in \mathcal{D}_n$. Si $x = 0$, alors $-x = 0 \in \mathcal{D}_n$. Si $x \neq 0$, alors il existe $f \in \mathfrak{a}$ de degré n tel que $x = \delta(f)$. Mais alors, $-x = -\delta(f) = \delta(-f)$. Comme $-f$ est un élément de $\mathfrak{a}$ de degré n, on obtient encore $-x \in \mathcal{D}_n$.

Montrons maintenant que $x + y \in \mathcal{D}_n$. Si l'un des éléments x, y ou $x + y$ est nul, il n'y a rien à faire. On suppose donc que x, y et $x + y$ sont tous les trois non nuls. Par définition, il existe f, $g \in \mathfrak{a}$ de degré n tels que $x = \delta(f)$ et $y = \delta(g)$. Puisque $x + y$ est non nul, $f + g$ est alors un élément de $\mathfrak{a}$ de degré n, et l'on a $x + y = \delta(f + g) \in \mathcal{D}_n$.

Enfin, montrons que $ax \in \mathcal{D}_n$. Encore une fois, si $x = 0$ ou $ax = 0$, il n'y a rien à faire. On suppose donc que x et ax sont tous deux non nuls, et l'on écrit $x = \delta(f)$, où $f \in \mathfrak{a}$ est de degré n. Comme $ax \neq 0$, af est un élément de $\mathfrak{a}$ de degré n, et l'on a $ax = \delta(af) \in \mathcal{D}_n$.

Ainsi, $\mathcal{D}_n$ est bien un idéal de A.

Remarquons maintenant que l'on a $\mathcal{D}_n \subset \mathcal{D}_{n+1}$ pour tout $n \geqslant 0$. En effet, soit $x \in \mathcal{D}_n$. Si $x = 0$, il n'y a rien à faire. Si $x \neq 0$, on a $x = \delta(f)$, où le polynôme $f \in \mathfrak{a}$ est de degré n. Mais alors, Xf est un élément de $\mathfrak{a}$ de degré $n + 1$, et vérifiant $x = \delta(Xf)$, d'où l'inclusion souhaitée.

Puisque A est noethérien, il existe donc un entier $r \geqslant 0$ tel que

$$\mathcal{D}_n = \mathcal{D}_r, \quad \text{pour tout } n \geqslant r,$$

par la proposition XVIII-4.6. On en déduit alors aisément l'égalité

$$\bigcup_{n\geqslant 0} \mathcal{D}_n = \mathcal{D}_r.$$

Par ailleurs, l'idéal $\mathcal{D}_n$ est de type fini pour tout $n \in [\![0, r]\!]$, et il existe donc une famille finie de polynômes $f_{n,1}, \ldots, f_{n,j_n} \in \mathfrak{a}$ de degré n dont les coefficients dominants engendrent l'idéal $\mathcal{D}_n$.

Soit $\mathfrak{b}$ l'idéal engendré par la famille

$$f_{n,j}, \quad n \in [\![0, r]\!], \ j \in [\![1, j_n]\!].$$

Nous allons démontrer l'égalité $\mathfrak{a} = \mathfrak{b}$. Comme on a clairement $\mathfrak{b} \subset \mathfrak{a}$, il suffit de démontrer l'inclusion inverse. Nous allons procéder par récurrence sur le degré d'un polynôme $g \in \mathfrak{a}$.

On peut supposer g non nul. Supposons tout d'abord que $\deg(g) = 0$. Alors, g est constant, donc égal à son propre coefficient dominant. Ainsi, on a $g \in \mathcal{D}_0$, et g est donc une combinaison A-linéaire (donc $A[X]$-linéaire) de $f_{01}, \ldots, f_{0,j_0}$, et l'on a fini.

Supposons avoir montré que tout élément de $\mathfrak{a}$ de degré $\leqslant m$ est dans $\mathfrak{b}$ et soit $g \in \mathfrak{a}$ de degré $m+1$. Nous allons démontrer l'existence de $f \in \mathfrak{b}$, de degré $m+1$, tel que $\delta(f) = \delta(g)$.

Supposons tout d'abord que $m+1 \leqslant r$. Comme $\delta(g) \in \mathcal{D}_{m+1}$, on peut écrire

$$\delta(g) = a_1\delta(f_{m+1,1}) + \cdots + a_{j_{m+1}}\delta(f_{m+1,j_{m+1}}), a_i \in A.$$

Le polynôme $f = a_1 f_{m+1,1} + \cdots + a_{m+1} f_{m+1,j_{m+1}} \in \mathfrak{b}$ convient alors.

Supposons maintenant que $m+1 \geqslant r$. Alors, on a

$$\delta(g) \in \mathcal{D}_{n+1} \subset \bigcup_{m\geqslant 0} \mathcal{D}_m = \mathcal{D}_r.$$

On peut donc écrire

$$\delta(g) = a_1\delta(f_{r1}) + \cdots + a_{j_r}\delta(f_{rj_r}), \quad a_i \in A.$$

Cette fois, le polynôme $f = X^{m+1-r}(a_1 f_{r1} + \cdots + a_r f_{rj_r})$ convient.

Par construction de f, le polynôme $g - f$ est un élément de $\mathfrak{a}$ de degré $\leqslant m$. Par hypothèse de récurrence, $g - f \in \mathfrak{b}$, et par conséquent $g = f + (g-f) \in \mathfrak{b}$. Cela achève la récurrence, ainsi que la démonstration. □

2.3. Corollaire. *Soit A un anneau noethérien. Alors, pour tout $n \geqslant 0$, l'anneau $A[X_1, \ldots, X_n]$ est noethérien.*

Avant de poursuivre, nous avons besoin d'une définition.

2.4. Définition. Soit K un corps. On dit qu'une K-algèbre associative, commutative et unitaire $\mathscr{A}$ est *de type fini* s'il existe $\alpha_1, \ldots, \alpha_n \in \mathscr{A}$ tels que $\mathscr{A} = K[\alpha_1, \ldots, \alpha_n] = \{P(\alpha_1, \ldots, \alpha_n) \mid P \in K[X_1, \ldots, X_n]\}$.

On a alors la proposition suivante, fondamentale en géométrie algébrique.

2.5. Proposition. *Soit K un corps. Alors, toute K-algèbre associative, commutative et unitaire de type fini est un anneau noethérien.*

Démonstration. Soit $\mathscr{A}$ une K-algèbre associative, commutative et unitaire de type fini. Si $\mathscr{A} = K[\alpha_1, \ldots, \alpha_n]$, on a donc un morphisme d'anneaux surjectif

$$\begin{aligned} K[X_1, \ldots, X_n] &\longrightarrow \mathscr{A} \\ P &\longmapsto P(\alpha_1, \ldots, \alpha_n). \end{aligned}$$

Or, K est un anneau noethérien, et par le corollaire 2.3, $K[X_1, \ldots, X_n]$ est noethérien. Mais alors, le lemme XVIII-4.3 montre que $\mathscr{A}$ est un anneau noethérien. Cela achève la démonstration. □

3. Idéaux premiers de $A[X]$, A principal

Pour finir ce chapitre, nous allons nous intéresser aux idéaux premiers de l'anneau $A[X]$, lorsque A est un anneau principal.

3.1. Théorème. *Soit A un anneau principal. Alors, les idéaux premiers non nuls de $A[X]$ sont exactement les idéaux fournis par la liste suivante :*

(1) $\mathfrak{p}_\pi = (\pi)$, *où π est irréductible ;*

(2) $\mathfrak{p}_f = (f)$, *où $f \in A[X]$ irréductible non constant ;*

(3) $\mathfrak{m}_{\pi,f} = (\pi, f)$, *où $\pi \in A$ est irréductible, et $f \in A[X]$ est un polynôme unitaire, irréductible modulo π.*

De plus :

(i) l'idéal $\mathfrak{p}_\pi$ est premier, non maximal ;

(ii) si A possède une infinité d'éléments irréductibles deux à deux non associés, alors l'idéal $\mathfrak{p}_f$ est premier, non maximal ;

(iii) l'idéal $\mathfrak{m}_{\pi,f}$ est maximal.

Démonstration. Commençons par montrer que les idéaux de la liste ci-dessus sont bien premiers, ainsi que les propriétés (i)-(iii).

Soit $\pi \in A$ un élément irréductible. Si $g \in A[X]$, on notera $\overline{g} \in (A/(\pi))[X]$ sa réduction modulo π. La réduction modulo π induit un isomorphisme

d'anneaux
$$A[X]/(\pi) \simeq \big(A/(\pi)\big)[X].$$

Puisque A est principal, tout élément irréductible engendre un idéal maximal, et $A/(\pi)$ est donc un corps. Par suite, $\big(A/(\pi)\big)[X]$ est un anneau intègre qui n'est pas un corps (puisque les seuls éléments inversibles sont les polynômes constants), et (π) est donc un idéal premier non maximal.

Soit maintenant f un polynôme irréductible non constant de $A[X]$. Comme l'anneau A est principal, donc factoriel par le théorème XVIII-5.16, $A[X]$ est factoriel par le théorème de Gauss. Ainsi, tout élément irréductible est premier par le lemme d'Euclide, d'où la première partie. Si de plus l'anneau A possède une infinité d'éléments irréductibles deux à deux non associés, il existe un élément irréductible π ne divisant pas le coefficient dominant de f. On a alors $(f) \subset (\pi, f)$, et l'inclusion est stricte. En effet, dans le cas contraire, on a aurait $\pi \in (f)$, c'est-à-dire $\pi = fg$, $g \in A[X]$, d'où une contradiction en comparant les degrés. Il faut maintenant vérifier que $(\pi, f) \neq A[X]$.

Supposons que $(\pi, f) = A[X]$. Il existe alors $g, h \in A[X]$ tels que $\pi g + fh = 1$. On a alors $\overline{f}\,\overline{h} = \overline{1} \in \big(A/(\pi)\big)[X]$. Ainsi, $\overline{f}$ est inversible dans $\big(A/(\pi)\big)[X]$. Comme (π) est un idéal premier, $A/(\pi)$ est intègre. On en déduit alors que $\deg(\overline{f}) = 0$. Or, π ne divisant pas le coefficient dominant de f, on obtient
$$\deg(f) = \deg(\overline{f}) = 0,$$

d'où une contradiction. Ainsi, (f) n'est pas maximal dans ce cas.

Enfin, soit π un élément irréductible de A, et soit f un polynôme unitaire de $A[X]$, irréductible modulo π.

Soit $\varphi : A[X] \longrightarrow \big(A/(\pi)\big)[X]/(\overline{f})$ le morphisme d'anneaux obtenu en composant le morphisme $A[X] \longrightarrow \big(A/(\pi)\big)[X]$ de réduction modulo π avec la projection canonique $A/(\pi)[X] \longrightarrow \big(A/(\pi)\big)[X]/(\overline{f})$.

Ce morphisme est clairement surjectif, de noyau (π, f), et induit donc un isomorphisme d'anneaux
$$A[X]/(\pi, f) \simeq \big(A/(\pi)\big)[X]/(\overline{f}).$$

Comme A est principal, $A/(\pi)$ est un corps, et $\big(A/(\pi)\big)[X]$ est donc un anneau principal. Comme $\overline{f}$ est irréductible, $\big(A/(\pi)\big)[X]/(\overline{f})$ est un corps, et (π, f) est maximal.

Il reste à démontrer que tout idéal premier non nul est bien d'une des trois formes précédentes. Soit $\mathfrak{p}$ un idéal premier non nul de $A[X]$. Alors, $\mathfrak{p} \cap A$ est un idéal premier de A. Comme A est principal, $\mathfrak{p} \cap A$ est soit nul, soit engendré par un élément irréductible $\pi \in A$.

Supposons tout d'abord que $\mathfrak{p} \cap A = (\pi)$, π irréductible. On va alors montrer que $\mathfrak{p}$ est de la forme $\mathfrak{p}_\pi$ ou $\mathfrak{m}_{\pi,f}$. Soit $\overline{\mathfrak{p}}$ l'image de $\mathfrak{p}$ par la réduction modulo π. On vérifie facilement que c'est un idéal premier de $\big(A/(\pi)\big)[X]$. Comme déjà remarqué précédemment, $A/(\pi)$ est un corps et $\big(A/(\pi)\big)[X]$ est principal. Ainsi, $\overline{\mathfrak{p}}$ est soit nul, soit engendré par un polynôme irréductible unitaire de $\big(A/(\pi)\big)[X]$. Dans le premier cas, on a donc $\mathfrak{p} \subset (\pi)$, et comme $\pi \in \mathfrak{p}$ par définition, on a donc $\mathfrak{p} = (\pi) = \mathfrak{p}_\pi$.

Dans le second cas, on a $\overline{\mathfrak{p}} = (\overline{f})$, où f est un polynôme unitaire de $A[X]$ tel que $\overline{f}$ soit irréductible. En particulier, il existe $g \in \mathfrak{p}$ tel que $\overline{f} = \overline{g}$. Mais alors, on peut écrire $f = g + \pi h$ avec $h \in A[X]$, et comme $\pi \in \mathfrak{p}$, on a également $f \in \mathfrak{p}$. Par conséquent, $\mathfrak{m}_{\pi,f} \subset \mathfrak{p}$. Comme $\mathfrak{p} \neq A$ puisque $\mathfrak{p}$ est premier, on déduit de la maximalité de $\mathfrak{m}_{\pi,f}$ que $\mathfrak{p} = \mathfrak{m}_{\pi,f}$.

Supposons maintenant que $\mathfrak{p} \cap A = (0)$. Soit f un élément non nul de $\mathfrak{p}$ de degré minimal. Ce degré est au moins égal à 1 par hypothèse. Écrivons maintenant $f = \alpha f_0$, avec $\alpha \in A \setminus \{0\}$ et $f_0 \in A[X]$ primitif. Comme $\mathfrak{p}$ est premier et ne contient pas d'élément non nul de A par hypothèse, on en déduit que $f_0 \in \mathfrak{p}$. Quitte à remplacer f par f_0, on peut donc supposer f primitif. Soit $g \in \mathfrak{p} \setminus \{0\}$. Par définition de f, on a $\deg(g) \geqslant \deg(f)$. On peut donc écrire

$$g = fh + r, \quad h, r \in K_A[X],\ \deg(r) < \deg(f).$$

Notons que h est non nul, car sinon, on aurait $g = r$ et $\deg(g) < \deg(f)$. Si $d \in A$ est un dénominateur commun des coefficients de h et de r, on a donc dh, $dr \in A[X]$. Mais alors, $dr = dg - f(dh) \in \mathfrak{p}$. Comme on a

$$\deg(dr) = \deg(r) < \deg(f),$$

on en déduit que $dr = 0$ par minimalité du degré de f, et donc $r = 0$. Ainsi, on obtient $g = fh$. D'après le lemme XXI-3.7, et puisque f est primitif, on a $c(g) = c(h)$. Puisque $g \in A[X]$, on en déduit que $c(h) \in A \setminus \{0\}$. En particulier, $h \in A[X]$, et l'on a finalement $\mathfrak{p} = (f) = \mathfrak{p}_f$, ce qui achève la démonstration. □

3.2. Corollaire. *Soit A un anneau principal possédant une infinité d'éléments irréductibles non associés deux à deux. Alors, les idéaux maximaux de $A[X]$ sont les idéaux de la forme (π, f), où $\pi \in A$ est un élément irréductible et $f \in A[X]$ est un polynôme unitaire irréductible modulo π.*

Démonstration. Remarquons tout d'abord qu'un idéal maximal de $A[X]$ est premier et non nul, puisque $A[X]$ n'est pas un corps. Il suffit alors d'appliquer le théorème précédent pour conclure. □

Le théorème précédent est bien entendu faux si A n'est pas principal. De plus, le corollaire précédent est également faux si A n'a pas assez d'éléments irréductibles.

3.3. Exemples

(1) Soit $A = \mathbb{C}[Y, Z]$. Alors, $\pi = Y$ est irréductible dans A, $f = X + Z$ est un élément irréductible de l'anneau $A/(Y)[X] \simeq \mathbb{C}[Z][X]$. En revanche, l'idéal $\mathfrak{m}_{\pi,f} = (Y, X + Z)$ n'est pas maximal, puisqu'il est contenu strictement dans l'idéal (X, Y, Z), qui est distinct de $A[X]$ (puisque ses éléments sont sans terme constant).

(2) Soit $A = \{\frac{r}{s} \in \mathbb{Q} \mid s \geqslant 1 \text{ impair}\}$. On vérifie aisément que A est un sous-anneau de $\mathbb{Q}$.

Montrons que A est principal. L'anneau A est intègre, comme sous-anneau de $\mathbb{Q}$. Soit $\mathfrak{a}$ un idéal de A. Alors, $\mathfrak{a} \cap \mathbb{Z}$ est un idéal de $\mathbb{Z}$, donc engendré par $r_0 \in \mathbb{Z}$. Par hypothèse, $r_0 \in \mathfrak{a}$, et ainsi $(r_0) \subset \mathfrak{a}$. Inversement, soit $\frac{r}{s} \in \mathfrak{a}$, avec $s \geqslant 1$ impair. Alors, on a

$$r = \frac{r}{s} + \cdots + \frac{r}{s} \in \mathfrak{a} \cap \mathbb{Z},$$

et par conséquent, il existe $m \in \mathbb{Z}$ tel que $r = r_0 m$. Mais alors,

$$\frac{r}{s} = r_0 \frac{m}{s} \in (r_0).$$

Ainsi, $\mathfrak{a} = (r_0)$, et A est un anneau principal.

Soit $\mathfrak{m} = (2X - 1)$. Montrons que $\mathfrak{m}$ est un idéal maximal de $A[X]$, dont l'intersection avec A est triviale. Ce dernier point étant clair par des considérations de degré, il reste à voir que $\mathfrak{m}$ est maximal. Remarquons que le morphisme d'évaluation

$$\begin{aligned} \varphi : A[X] &\longrightarrow \mathbb{Q} \\ f &\longmapsto f\big(\tfrac{1}{2}\big) \end{aligned}$$

est surjectif.

En effet, si $x \in \mathbb{Q}$, alors il existe un entier $m \geqslant 0$ tel que $2^m x = a \in A$. Mais alors, $\varphi(aX^m) = x$.

Montrons maintenant que $\mathrm{Ker}(\varphi) = \mathfrak{m}$. Soit $f \in \mathrm{Ker}(\varphi)$. En considérant f comme un élément de $\mathbb{Q}[X]$, on voit que

$$f = (2X - 1)g, \quad g \in \mathbb{Q}[X].$$

Par le lemme XXI-3.7, on a alors

$$c(f) = c\big((2X - 1)g\big) = c(2X - 1)c(g) = c(g).$$

Comme $f \in A[X]$, on a $c(g) = c(f) \in A$, d'où $g \in A[X]$ d'après le lemme XXI-3.6 (2). Ainsi, $\mathrm{Ker}(\varphi) \subset (2X - 1)$. L'autre inclusion étant évidente, on a $\mathrm{Ker}(\varphi) = \mathfrak{m}$, et l'on obtient dès lors

$$A[X]/\mathfrak{m} \simeq \mathbb{Q}.$$

Ainsi, $\mathfrak{m}$ est maximal.

Comme $\mathfrak{m}$ est maximal, $\mathfrak{m}$ est premier. Comme $A[X]$ est intègre, cela montre que $2X-1$ est un polynôme irréductible de $A[X]$ non constant (cf. lemme XVIII-1.19) qui engendre un idéal maximal.
Cela ne contredit nullement le théorème 3.1, puisque l'on peut montrer que le seul élément irréductible de A est 2, à association près.

4. Exercices

E1. Soit A un anneau principal, et soit $f \in A[X]$ un polynôme irréductible non constant. Montrer que (f) est un idéal maximal si, et seulement si, le coefficient dominant de f est divisible par tous les éléments irréductibles de A.

E2. Soit K un corps.

a. On suppose que $K = \mathbb{C}$. Montrer que tout idéal maximal de $K[X,Y]$ est la forme $(X-a, Y-b)$, $a, b \in \mathbb{C}$.

b. On suppose que $K = \mathbb{R}$. Montrer que tout idéal maximal de $K[X,Y]$ est d'une des trois formes suivantes :
(i) $(X-a, Y-b)$, $a, b \in \mathbb{R}$;
(ii) $(\pi(X), Y+aX+b)$, avec $\pi \in \mathbb{R}[X]$ irréductible unitaire, et $a, b \in \mathbb{R}$;
(iii) $\big(X-a, \pi(Y)\big)$, avec $\pi \in \mathbb{R}[X]$ irréductible unitaire de degré 2.

E3. Soit $A = \mathbb{R}[X,Y]/(X^2+Y^2+1)$.

a. Montrer que A est intègre et noethérien.

b. Montrer que tout élément de A s'écrit de manière unique sous la forme

$$\overline{R(X)+YS(X)}, \quad R, S \in \mathbb{R}[X].$$

c. Montrer que $A^\times = \{\overline{\lambda} \mid \lambda \in \mathbb{R}^\times\}$.

d. Montrer que pour tout $\overline{P} \in A \setminus A^\times$, l'application

$$\begin{aligned} \mathbb{R} &\longrightarrow A/(\overline{P}) \\ \lambda &\longmapsto \overline{\lambda} \end{aligned}$$

n'est pas surjective. En utilisant l'exercice XVIII-16, en déduire que l'anneau A n'est pas euclidien.

e. Utiliser l'exercice précédent pour montrer qu'un idéal maximal $\mathfrak{m}$ de l'anneau $\mathbb{R}[X,Y]$ contenant X^2+Y+1 est de la forme

$$\mathfrak{m} = (X^2+Y^2+1, aX+bY+c), \quad a, b, c \in \mathbb{R}.$$

En déduire que tout idéal maximal de A est principal.

f. Montrer que A est un anneau principal, non euclidien.

Chapitre XXIII

Séries formelles

Dans ce chapitre, nous allons étudier l'anneau des séries formelles sur un anneau commutatif A, brièvement introduit dans le chapitre XV.

1. Premières propriétés

Soit A un anneau. On rappelle (cf. définition XV-6.2) que l'anneau $A[[X]]$ des séries formelles à coefficients dans A est l'ensemble des sommes formelles

$$\sum_{n\geqslant 0} a_n X^n, \quad a_n \in A,$$

muni de la somme

$$\sum_{n\geqslant 0} a_n X^n + \sum_{n\geqslant 0} b_n X^n = \sum_{n\geqslant 0} (a_n + b_n) X^n$$

et du produit

$$\Big(\sum_{n\geqslant 0} a_n X^n\Big)\Big(\sum_{n\geqslant 0} b_n X^n\Big) = \sum_{n\geqslant 0} \Big(\sum_{k=0}^{n} a_k b_{n-k}\Big) X^n,$$

les éléments neutres étant respectivement les séries formelles constantes 0 et 1.

Si $S = \sum_{n\geqslant 0} a_n X^n \in A[[X]]$, le coefficient $a_0 \in A$ figurant dans l'expression de S est appelé le *terme constant* de S.

On se restreindra dans ce chapitre au cas où A est un anneau commutatif. En particulier, $A[[X]]$ est un anneau commutatif. On commence par étudier l'intégrité de l'anneau $A[[X]]$.

Contrairement au cas des polynômes, nous n'avons pas à notre disposition une notion de degré, notion pourtant fort utile pour étudier l'intégrité d'un anneau de polynômes. En revanche, nous avons la notion d'ordre d'une série formelle.

1.1. Définition. Soit $S = \sum_{n\geqslant 0} a_n X^n \in A[[X]]$ non nulle.

L'ordre de S, noté $v_X(S)$, est le plus petit entier $m \geqslant 0$ tel que $a_m \neq 0$. On pose également $v_X(0) = +\infty$.

Le lemme suivant résume les propriétés de l'ordre d'une série formelle.

1.2. Lemme. *Soit A un anneau commutatif. Alors :*

(1) *pour tout $S \in A[[X]]$, on a $v_X(S) = +\infty$ si, et seulement si, $S = 0$;*

(2) *pour tout $S \in A[[X]]$, on a $v_X(-S) = v_X(S)$;*

(3) *pour tous $S, T \in A[[X]]$, on a $v_X(S+T) \geqslant \min\big(v_X(S), v_X(T)\big)$, avec égalité si $v_X(S) \neq v_X(T)$;*

(4) *pour tous $S, T \in A[[X]]$, on a $v_X(ST) \geqslant v_X(S) + v_X(T)$. Si de plus A est intègre, alors on a*

$$v_X(ST) = v_X(S) + v_X(T), \quad \text{pour tous } S,\, T \in A[[X]].$$

Démonstration. Les points (1) et (2) proviennent simplement de la définition. Les points (3) et (4) étant clairs si $S = 0$ ou $T = 0$, on peut supposer que S et T sont non nuls. Écrivons $S = \sum_{k\geqslant 0} a_k X^k$ et $T = \sum_{k\geqslant 0} b_k X^k$. Ainsi, on a $S + T = \sum_{k\geqslant 0} (a_k + b_k) X^k$.

Posons $n = v_X(S)$ et $m = v_X(T)$. On peut toujours supposer que $n \leqslant m$. On a alors $a_k = b_k = 0$ pour tout entier $k \in [\![0, n-1]\!]$, puisque $n \leqslant m$. Mais alors, $a_k + b_k = 0$ pour tout entier $k \in [\![0, n-1]\!]$, et donc $v_X(S+T) \geqslant n$.

Supposons maintenant que $n < m$. On a donc $b_n = 0$, et par conséquent, on a $a_n + b_n = a_n \neq 0$, ce qui démontre que $v_X(S+T) = n$ dans ce cas, d'où (3).

Démontrons le point (4). On a $S = X^n \sum_{k\geqslant 0} a_{k+n} X^k$ et $T = X^m \sum_{k\geqslant 0} b_{k+m} X^k$, et ainsi

$$\begin{aligned} ST &= X^{n+m} \sum_{k\geqslant 0} \Big(\sum_{i=0}^{k} a_{i+n} b_{k-i+m} \Big) X^k \\ &= \sum_{k\geqslant 0} \Big(\sum_{i=0}^{k} a_{i+n} b_{k-i+m} \Big) X^{k+n+m} \\ &= \sum_{k\geqslant n+m} \Big(\sum_{i=0}^{k-n-m} a_{i+n} b_{k-i-n} \Big) X^k. \end{aligned}$$

Cela montre que le premier terme de ST qui soit éventuellement non nul est $a_m b_n X^{n+m}$. On a donc $v_X(ST) \geqslant n+m$. Si de plus l'anneau A est intègre, on a alors $a_m b_n \neq 0$ puisque a_m et b_n sont non nuls, et dans ce cas, on a donc $v_X(ST) = n+m$. Cela achève la démonstration. □

1.3. Corollaire. *Soit A un anneau commutatif. Alors, $A[[X]]$ est intègre si, et seulement si, A est intègre.*

Démonstration. Supposons que l'anneau $A[[X]]]$ soit intègre. Puisque A s'identifie à un sous-anneau de $A[[X]]$, il est alors intègre. Inversement, supposons que A soit intègre. Alors, $A[[X]]$ est clairement commutatif et non trivial. Soient $S, T \in A[[X]]$ deux séries formelles non nulles. D'après le point (4) du lemme précédent, on a $v_X(ST) = v_X(S) + v_X(T)$. En particulier, $v_X(ST) \neq +\infty$, et $ST \neq 0$ d'après le point (1) de ce même lemme. Cela achève la démonstration. □

Déterminons maintenant les éléments inversibles de $A[[X]]$.

1.4. Lemme. *Soit A un anneau commutatif. Alors, $S \in A[[X]]$ est inversible si, et seulement si, son terme constant est inversible dans A.*

Démonstration. La série formelle $S = \sum_{n\geqslant 0} a_n X^n$ est inversible si, et seulement si, il existe $T = \sum_{n\geqslant 0} b_n X^n \in A[[X]]$ telle que $ST = 1$. Or, on a

$$ST = \sum_{n\geqslant 0} \Big(\sum_{k=0}^{n} a_k b_{n-k} \Big) X^n = a_0 b_0 + \sum_{n\geqslant 1} \Big(\sum_{k=0}^{n} a_k b_{n-k} \Big) X^n.$$

Ainsi, S est inversible si, et seulement si, il existe une suite $(b_n)_{n\geqslant 0}$ d'éléments de A telle que

$$a_0 b_0 = 1 \quad \text{et} \quad \sum_{k=0}^{n} a_k b_{n-k} = 0, \quad \text{pour tout } n \geqslant 1.$$

En particulier, si S est inversible, alors il existe $b_0 \in A$ tel que $a_0 b_0 = 1$, et a_0 est donc inversible dans A.

Inversement, supposons que $a_0 \in A^\times$. On définit alors par récurrence une suite $(b_n)_{n\geqslant 0}$ d'éléments de A en posant

$$b_0 = a_0^{-1} \quad \text{et} \quad b_n = -a_0^{-1} \Big(\sum_{k=1}^{n} a_k b_{n-k} \Big), \quad \text{pour tout } n \geqslant 1.$$

Par construction, la série formelle $T = \sum_{n\geqslant 0} b_n X^n$ est alors un inverse de S, d'où le résultat. □

Notation. Si $S \in A[[X]]^\times$, son inverse sera noté S^{-1}.

1.5. Exemple. Le polynôme $1 - X \in A[X]$ est inversible dans $A[[X]]$, puisque 1 est inversible. De plus, on vérifie facilement que

$$(1 - X) \sum_{n \geqslant 0} X^n = 1,$$

et par suite

$$(1 - X)^{-1} = \sum_{n \geqslant 0} X^n.$$

1.6. Proposition. *Soit K un corps. Alors, $K[[X]]$ est intègre. De plus :*

(1) *pour toute série formelle $S \in K[[X]]$, on a $S \in K[[X]]^\times$ si, et seulement si, $v_X(S) = 0$;*

(2) *toute série formelle $S \in K[[X]]$ non nulle s'écrit de manière unique sous la forme*

$$S = X^m U, \quad m \geqslant 0,\ U \in K[[X]]^\times.$$

De plus, $m = v_X(S)$;

(3) *tout idéal non nul de $K[[X]]$ est de la forme (X^m) pour un unique entier $m \geqslant 0$. De plus, (X) est l'unique idéal maximal de $K[[X]]$. En particulier, $K[[X]]$ est principal.*

Démonstration. Un corps étant intègre, l'anneau $K[[X]]$ est intègre par le corollaire 1.3. D'après le lemme 1.4, $S = \sum_{n \geqslant 0} a_n X^n \in K[[X]]$ est inversible si, et seulement si, a_0 est inversible. Comme K est un corps, cela revient à dire que $a_0 \neq 0$, c'est-à-dire $v_X(S) = 0$.

Soit $S = \sum_{n \geqslant 0} a_n X^n$ une série formelle non nulle. Soit $m = v_X(S)$. On a alors

$$S = \sum_{n \geqslant m} a_n X^n = X^m \sum_{n \geqslant 0} a_n X^{n-m} = X^m U,$$

avec $U = \sum_{n \geqslant 0} a_{n+m} X^n$. Puisque le terme constant de U est a_m, qui est non nul, U est inversible, d'où l'existence de la décomposition annoncée.

Supposons que $S = X^m U = X^n V$, avec $n, m \geqslant 0$ et $U, V \in K[[X]]^\times$. En particulier, $v_X(U) = v_X(V) = 0$. D'après le lemme 1.2, on a alors

$$v_X(S) = v_X(X^m) + v_X(U) = m + 0 = m,$$

et de même

$$v_X(S) = v_X(X^n) + v_X(V) = n + 0 = n.$$

Ainsi, $m = n = v_X(S)$. Mais alors, on a $X^m(U - V) = 0$, et par intégrité, on obtient $U - V = 0$, soit $U = V$. Cela montre l'unicité de la décomposition.

Il reste à démontrer le dernier point. Soit I un idéal non nul de $K[[X]]$, et soit $S \in I \setminus \{0\}$ telle que $v_X(S)$ soit minimale. D'après (2), on a $S = X^m U$,

avec $U \in K[[X]]^\times$ et $m = v_X(S)$. Montrons que $I = (X^m)$. Puisque $S \in I$, on a $(S) \subset I$. Or, $(S) = (X^m U) = (X^m)$, car U est inversible, d'où l'inclusion $(X^m) \subset I$. Si maintenant $T \in I$, alors $v_X(T) \geqslant m$ par définition de m. On a donc $T = \sum_{k \geqslant n} a_k X^k$, avec $n \geqslant m$. Mais alors,

$$T = X^m \Big(\sum_{k \geqslant n} a_k X^{k-m} \Big) = X^m \Big(\sum_{k \geqslant n-m} a_{k+m} X^k \Big) \in (X^m),$$

et donc $I \subset (X^m)$. Finalement, on a $I = (X^m)$.

Si maintenant $I = (X^m) = (X^n)$, on a $X^m = X^n S_0$, avec $S_0 \in K[[X]]^\times$, et ainsi

$$m = v_X(X^m) = n + v_X(S_0) = n + 0 = n.$$

Ce qui précède montre en particulier que tout idéal propre de $K[[X]]$ est contenu dans (X), qui est lui-même un idéal distinct de $K[[X]]$. Ainsi, (X) est un idéal maximal. De plus, si $\mathfrak{m}$ est un idéal maximal, on a $\mathfrak{m} \subset (X)$, et par maximalité de $\mathfrak{m}$, on a $\mathfrak{m} = (X)$. Cela achève la démonstration. □

Nous continuons ce paragraphe par quelques considérations sur les séries formelles nilpotentes.

1.7. Lemme. *Soit A un anneau commutatif. Alors,*

$$\mathrm{Nil}(A[[X]]) \subset \mathrm{Nil}(A)[[X]] = \Big\{ \sum_{n \geqslant 0} a_n X^n \mid a_n \in \mathrm{Nil}(A), \quad \textit{pour tout } n \geqslant 0 \Big\},$$

avec égalité dès que $\mathrm{Nil}(A)$ *est de type fini.*

Démonstration. Pour tout $n \geqslant 0$, soit (H_n) la propriété suivante :

(H_n) Pour tout $S = \sum_{n \geqslant 0} a_n X^n \in \mathrm{Nil}(A[[X]])$, on a $a_k \in \mathrm{Nil}(A)$, pour tout entier $k \in [\![0, n]\!]$.

Soit $S = \sum_{n \geqslant 0} a_n X^n \in \mathrm{Nil}(A[[X]])$, et soit $m \geqslant 1$ tel que $S^m = 0$. Le terme constant de S^m est a_0^m, et par conséquent $a_0^m = 0$. Ainsi, $a_0 \in \mathrm{Nil}(A)$ et la propriété (H_0) est vraie.

Soit $k \geqslant 0$, et supposons que (H_k) soit vraie. Montrons que (H_{k+1}) est vraie. Par hypothèse de récurrence, $a_0, \ldots, a_k \in \mathrm{Nil}(A)$. Il reste donc à démontrer que $a_{k+1} \in \mathrm{Nil}(A)$. Puisque $\mathrm{Nil}(A[[X]])$ est un idéal, on a

$$S - (a_0 + a_1 X + \cdots + a_k X^k) \in \mathrm{Nil}(A),$$

c'est-à-dire $\sum_{n \geqslant k+1} a_n X^n = X^{k+1} \sum_{n \geqslant 0} a_{n+k+1} X^n \in \mathrm{Nil}(A[[X]])$. Il existe donc $\ell \geqslant 1$ tel que

$$X^{\ell(k+1)} \Big(\sum_{n \geqslant 0} a_{n+k+1} X^n \Big)^\ell = 0.$$

Or, la multiplication par X dans $A[[X]]$ est clairement injective. On en déduit immédiatement que $\big(\sum_{n\geqslant 0} a_{n+k+1}X^n\big)^\ell = 0$. Ainsi, $\sum_{n\geqslant 0} a_{n+k+1}X^n$ est nilpotente. Comme (H_0) est vraie, on en déduit que le terme constant de cette série formelle est nilpotent, c'est-à-dire $a_{n+1} \in \mathrm{Nil}(A)$.

Ainsi, (H_{n+1}) est vraie. Le principe de récurrence montre que (H_n) est vraie pour tout $n \geqslant 0$.

Supposons maintenant que $\mathrm{Nil}(A)$ soit de type fini. Nous allons montrer qu'il existe $m \geqslant 1$ tel que

$$x_1 \cdots x_m = 0, \quad \text{pour tous } x_1, \ldots, x_m \in \mathrm{Nil}(A).$$

Supposons ce fait acquis, et soit $S = \sum_{n\geqslant 0} a_n X^n$, avec $a_n \in \mathrm{Nil}(A)$ pour tout entier $n \geqslant 0$. On a alors

$$S^m = \sum_{n\geqslant 0} \Big(\sum_{i_1+\cdots+i_m=n} a_{i_1} \cdots a_{i_m} \Big) X^n = 0,$$

d'où le résultat.

Il reste à démontrer le fait susmentionné. Écrivons $\mathrm{Nil}(A) = (\nu_1, \ldots, \nu_r)$, où $r \geqslant 0$. Soit $\ell \geqslant 1$ tel que $\nu_1^\ell = \cdots = \nu_r^\ell = 0$, et soit $m = \ell r$.

Tout élément de $\mathrm{Nil}(A)$ est de la forme $\alpha_1\nu_1 + \cdots + \alpha_r\nu_r, \alpha_i \in A$, et ainsi un produit de m éléments de $\mathrm{Nil}(A)$ est une combinaison A-linéaire de produits $\nu_1^{s_1} \cdots \nu_r^{s_r}$, où $s_1, \ldots, s_r \geqslant 0$ vérifient $s_1 + \cdots + s_r = m = \ell r$. Au moins un des s_i est supérieur ou égal à ℓ, car dans le cas contraire, on aurait $s_1 + \cdots + s_r < m$. Mais alors, $\nu_1^{s_1} \cdots \nu_r^{s_r} = 0$, d'où le fait annoncé. Cela achève la démonstration. □

1.8. Remarque. En particulier, l'égalité $\mathrm{Nil}(A[[X]]) = \mathrm{Nil}(A)[[X]]$ est vraie lorsque A est noethérien.

La dernière partie du lemme 1.7 est fausse sans hypothèse sur A, comme le montre l'exemple suivant.

1.9. Exemple. Soit $A = \mathbb{C}[X_0, X_1, \ldots]/(X_0^2, X_1^2, \ldots)$.

Pour tout $n \geqslant 0$, soit $a_n = \overline{X}_n$. On a donc $a_n^2 = 0$, pour tout $n \geqslant 0$. En particulier, $a_n \in \mathrm{Nil}(A)$, pour tout $n \geqslant 0$.

Soit $S = \sum_{n\geqslant 0} a_n X^n$ et soit $m \geqslant 1$. Écrivons $S^m = \sum_{n\geqslant 0} b_n X^n$. On a alors

$$b_{\frac{m(m-1)}{2}} = \sum_{i_1+\cdots+i_m=\frac{m(m-1)}{2}} a_{i_1} \cdots a_{i_m}.$$

Les seuls termes non nuls dans cette somme sont ceux qui pour lesquels les indices $i_1, \ldots, i_m$ sont deux à deux distincts. Remarquons maintenant que si $0 \leqslant j_1 < \cdots < j_m$ sont des entiers vérifiant

$$j_1 + \cdots + j_m = \frac{m(m-1)}{2},$$

on a $j_k \geqslant k-1$ pour tout $k \in [\![1, m]\!]$. Mais alors, $j_1 + \cdots + j_m \geqslant \dfrac{m(m-1)}{2}$. On en déduit aisément que $j_k = k-1$ pour tout $k \in [\![1, m]\!]$.

Cela démontre que $b_{\frac{m(m-1)}{2}} = ca_0 \cdots a_{m-1}$, pour un certain entier $c \geqslant 1$.

Nous allons vérifier que cet élément est non nul, et donc que $S^m \neq 0$. Supposons que $ca_0 \cdots a_{m-1} = 0$. Alors, il existe deux entiers k, $s \geqslant 0$ et des polynômes $P_0, \ldots, P_k \in \mathbb{C}[X_0, X_1, \ldots, X_s]$ tels que

$$cX_0 \cdots X_{m-1} = P_0 X_1^2 + \cdots + P_k X_k^2.$$

Cela est impossible, car $c \neq 0$ et le polynôme $P_0 X_1^2 + \cdots + P_k X_k^2$ n'a pas de terme en $X_0 \cdots X_{m-1}$.

Ainsi, $S^m \neq 0$ pour tout $m \geqslant 1$, et S n'est pas nilpotente.

1.10. Remarque. Le lecteur intéressé par des résultats plus précis sur les séries formelles nilpotentes pourra consulter [28]. Dans cet article, on peut également trouver une caractérisation des diviseurs de zéro de $A[[X]]$ lorsque A est noethérien, similaire à celle donnée pour les anneaux de polynômes par la proposition XV-6.33.

Nous allons maintenant nous intéresser à la factorialité de $A[[X]]$. On peut démontrer qu'il existe des anneaux factoriels pour lesquels $A[[X]]$ n'est pas factoriel, mais la preuve de ce fait est extrêmement délicate et sophistiquée (cf. [41, Théorème 4.1]).

En revanche, tout se passe mieux si A est principal.

1.11. Théorème. *Soit A un anneau principal. Alors, $A[[X]]$ est factoriel.*

Démonstration. Un anneau principal étant intègre, $A[[X]]$ l'est également par le corollaire 1.3. Nous allons montrer que tout idéal premier non nul contient un élément premier, ce qui permettra de conclure grâce au théorème XVIII-5.17.

Remarquons que le morphisme d'anneaux d'évaluation en 0

$$\begin{aligned} ev_0 : A[[X]] &\longrightarrow A \\ \sum_{n \geqslant 0} a_n X^n &\longmapsto a_0 \end{aligned}$$

est surjectif, de noyau (X), d'où un isomorphisme d'anneaux

$$A[[X]]/(X) \simeq A.$$

L'anneau A étant intègre, (X) est un idéal premier non nul, i.e. X est un élément premier de $A[[X]]$.

Soit $\mathfrak{p}$ un idéal premier non nul. Si $X \in \mathfrak{p}$, on a fini. On peut donc supposer sans perte de généralité que $X \notin \mathfrak{p}$. Puisque ev_0 est surjectif, $\mathfrak{a} = ev_0(\mathfrak{p})$ est un idéal de A. Montrons que $\mathfrak{a}$ est non nul. Si $S \in \mathfrak{p} \setminus \{0\}$, on peut toujours écrire $S = X^m T$, avec $m \geqslant 0$ et T de terme constant non nul. Comme $X \notin \mathfrak{p}$ et que $\mathfrak{p}$ est premier, on en déduit que $T \in \mathfrak{p}$. Mais alors, le terme constant de T est un élément non nul de $\mathfrak{a}$.

Puisque A est principal, on peut donc écrire $\mathfrak{a} = (\alpha)$, avec $\alpha \in A \setminus \{0\}$. Soit alors $S_0 \in \mathfrak{p}$ dont le terme constant est égal à α. Nous allons vérifier l'égalité $\mathfrak{p} = (S_0)$. Soit $S \in \mathfrak{p}$. Nous allons construire une suite $(b_n)_{n\geqslant 0}$ d'éléments de A telle que

$$S - \Big(\sum_{k=0}^{m} b_k X^k\Big) S_0 \in (X^{m+1}), \quad \text{pour tout } m \geqslant 0.$$

Puisque $S \in \mathfrak{p}$, le terme constant de S est un multiple de α par définition de α et de $\mathfrak{a}$. Il existe donc $b_0 \in A$ et $T_0 \in A[[X]]$ tels que $S = \alpha b_0 + XT_0$. Mais alors, $S - b_0 S_0 = b_0(\alpha - S_0) + XT_0 \in (X)$, puisque le terme constant de $\alpha - S_0$ est nul.

Soit $m \geqslant 0$, et supposons maintenant que $b_0, \ldots, b_m$ soient construits. Il existe donc $T_m \in A[[X]]$ telle que

$$S - \Big(\sum_{k=0}^{m} b_k X^k\Big) S_0 = X^{m+1} T_m.$$

Or, S et S_0 sont des éléments de $\mathfrak{p}$, donc $X^{m+1} T_m \in \mathfrak{p}$. Comme $X \notin \mathfrak{p}$ et que $\mathfrak{p}$ est premier, on en déduit que $T_m \in \mathfrak{p}$. Comme précédemment, on peut donc écrire $T_m = \alpha b_{m+1} + XT_{m+1}$, avec $b_{m+1} \in A$ et $T_{m+1} \in A[[X]]$. Mais alors, on a $T_m - b_{m+1} S_0 = b_{m+1}(\alpha - S_0) + XT_{m+1} \in (X)$. On en déduit alors que

$$S - \Big(\sum_{k=0}^{m+1} b_k X^k\Big) S_0 = X^{m+1}(T_m - b_{m+1} S_0) \in (X^{m+2}),$$

d'où l'existence de la suite $(b_n)_{n\geqslant 0}$ avec les propriétés désirées.

Pour tout $m \geqslant 0$, on a alors

$$S - \Big(\sum_{n\geqslant 0} b_n X^n\Big) S_0 = \Big(S - \Big(\sum_{k=0}^{m} b_k X^k\Big) S_0\Big) + \Big(\sum_{n\geqslant m+1} b_n X^n\Big) S_0 \in (X^{m+1}).$$

Si l'on écrit $S - \big(\sum_{n\geqslant 0} b_n X^n\big) S_0 = \sum_{n\geqslant 0} a_n X^n$, cela signifie que, pour tout entier $m \geqslant 0$, on a $a_0 = \cdots = a_m = 0$. Autrement dit, on a $a_n = 0$ pour tout $n \geqslant 0$, soit

$$S = \Big(\sum_{n\geqslant 0} b_n X^n\Big) S_0.$$

Ainsi, on a bien l'égalité $\mathfrak{p} = (S_0)$. Comme $\mathfrak{p}$ est un idéal premier non nul, S_0 est un élément premier de $A[[X]]$, contenu dans $\mathfrak{p}$. Cela achève la démonstration. □

1.12. Remarque. On a en fait démontré un résultat un peu plus précis : si A est un anneau principal, tout idéal premier non nul de $A[[X]]$ ne contenant pas X est principal.

On peut également déterminer tous les idéaux premiers contenant X (cf. exercice 3).

2. Composition et dérivation de séries formelles

Nous allons maintenant définir deux nouvelles opérations sur les séries formelles : la dérivation et la composition.

Nous commençons par la notion de famille sommable.

2.1. Définition. Soit $(S_i)_{i\in I}$ une famille de séries formelles de $A[[X]]$. Pour tout $i \in I$, on note

$$S_i = \sum_{n\geqslant 0} a_{n,i} X^n.$$

On dit que la famille $(S_i)_{i\in I}$ est *sommable* si pour tout entier $n \geqslant 0$, l'ensemble $\{i \in I \mid a_{n,i} \neq 0\}$ est fini.

Dans ce cas, on pose

$$\sum_{i\in I} S_i = \sum_{n\geqslant 0} \Big(\sum_{i\in I} a_{n,i}\Big) X^n.$$

La série formelle $\sum_{i\in I} S_i$ est alors appelée la *somme* de la famille $(S_i)_{i\in I}$.

2.2. Exemples

(1) Pour toute suite $(a_n)_{n\geqslant 0}$ d'éléments de A, la famille $(a_n X^n)_{n\geqslant 0}$ est évidemment sommable, et sa somme n'est rien d'autre que la série formelle $\sum_{n\geqslant 0} a_n X^n$.

(2) Soit $(S_k)_{k\geqslant 0}$ une suite de séries formelles. Si $v_X(S_k) \geqslant k$ pour tout entier $k \geqslant 0$, alors la famille $(S_k)_{k\geqslant 0}$ est sommable.

En effet, écrivons

$$S_k = \sum_{n\geqslant 0} a_{n,k} X^n.$$

Par hypothèse, pour tout $k \geqslant 0$, on a $a_{n,k} = 0$ pour tout $n \in [\![0, k-1]\!]$. Autrement dit, pour tout $n \geqslant 0$, on a $a_{n,k} = 0$, dès que $k > n$, et l'ensemble $\{k \in \mathbb{N} \mid a_{n,k} \neq 0\}$ est donc fini pour tout $n \geqslant 0$.

(3) Soit $S \in A[[X]]$ telle que $v_X(S) \geqslant 1$ et soit $(a_k)_{k\geqslant 0}$ une suite d'éléments de A. D'après le lemme 1.2, on a

$$v_X(a_k S^k) \geqslant v_X(a_k) + k v_X(S) \geqslant k, \qquad \text{pour tout } k \geqslant 0.$$

D'après l'exemple précédent, la famille $(a_k S^k)_{k\geqslant 0}$ est sommable.

2.3. Proposition. *Soit $(S_i)_{i\in I}$ une famille sommable d'éléments de $A[[X]]$.*

(1) *Pour tout $J \subset I$, la famille $(S_j)_{j\in J}$ est sommable.*

(2) *Pour tout $T \in A[[X]]$, la famille $(S_i T)_{i\in I}$ est sommable, et l'on a*

$$\sum_{i\in I} (S_i T) = \Big(\sum_{i\in I} S_i\Big) T.$$

(3) *Pour toute partition $(I_\lambda)_{\lambda\in\Lambda}$ de I, chaque famille $(S_i)_{i\in I_\lambda}$ est sommable, la famille $\big(\sum_{i\in I_\lambda} S_i\big)_{\lambda\in\Lambda}$ est sommable, et l'on a*

$$\sum_{i\in I} S_i = \sum_{\lambda\in\Lambda} \Big(\sum_{i\in I_\lambda} S_i\Big).$$

Démonstration. Le point (1) découle immédiatement de la définition d'une famille sommable. Pour tout $i \in I$, écrivons

$$S_i = \sum_{n\geqslant 0} a_{n,i} X^n.$$

On écrit aussi $T = \sum_{n\geqslant 0} b_n X^n$. On a alors

$$S_i T = \sum_{n\geqslant 0} \Big(\sum_{m=0}^{n} a_{m,i} b_{n-m} \Big) X^n.$$

Pour tout $i \in I$ et tout $n \geqslant 0$, on pose $c_{n,i} = \sum_{m=0}^{n} a_{m,i} b_{n-m}$.

Soit $n \geqslant 0$. Pour tout $m \geqslant 0$, l'ensemble $\{i \in I \mid a_{m,i} \neq 0\}$ est fini. En particulier, pour tout $n \geqslant 0$, l'ensemble

$$\bigcup_{m=0}^{n} \{i \in I \mid a_{m,i} \neq 0\}$$

est fini, et par suite l'ensemble $\{i \in I \mid c_{m,i} \neq 0\}$ est fini, puisqu'il est inclus dans l'ensemble $\bigcup_{m=0}^{n} \{i \in I \mid a_{m,i} \neq 0\}$. La famille $(S_i T)_{i \in I}$ est donc sommable. On a alors

$$\sum_{i \in I} (S_i T) = \sum_{n \geqslant 0} \sum_{i \in I} \Big(\sum_{m=0}^{n} a_{m,i} b_{n-m} \Big) X^n.$$

Par ailleurs, on a

$$\Big(\sum_{i \in I} S_i \Big) T = \Big(\sum_{n \geqslant 0} \Big(\sum_{i \in I} a_{n,i} \Big) X^n \Big) \Big(\sum_{n \geqslant 0} b_n X^n \Big),$$

ou encore

$$\Big(\sum_{i \in I} S_i \Big) T = \sum_{n \geqslant 0} \Big(\sum_{m=0}^{n} \Big(\sum_{i \in I} a_{m,i} \Big) b_{n-m} \Big) X^n.$$

Or, l'ensemble $\{i \in I \mid a_{m,i} \neq 0\}$ est fini, et par distributivité, on a

$$\Big(\sum_{i \in I} a_{m,i} \Big) b_{n-m} = \sum_{i \in I} a_{m,i} b_{n-m}.$$

Cette dernière somme étant finie, on a alors

$$\sum_{m=0}^{n} \Big(\sum_{i \in I} a_{m,i} b_{n-m} \Big) = \sum_{i \in I} \Big(\sum_{m=0}^{n} a_{m,i} b_{n-m} \Big),$$

d'où l'égalité désirée.

Il reste à démontrer (3). La sommabilité de la famille $(S_i)_{i \in I_\lambda}$ provient du point (1). Soit $\lambda \in \Lambda$. Par définition, on a

$$\sum_{i \in I_\lambda} S_i = \sum_{n \geqslant 0} \Big(\sum_{i \in I_\lambda} a_{n,i} \Big) X^n.$$

Soit $n \geqslant 0$. On doit démontrer que l'ensemble

$$\Big\{ \lambda \in \Lambda \mid \sum_{i \in I_\lambda} a_{n,i} \neq 0 \Big\}$$

est fini. Si $\lambda \in \Lambda$ est tel que $\sum_{i \in I_\lambda} a_{n,i} \neq 0$, alors, en particulier, il existe au moins un indice $i \in I_\lambda$ tel que $a_{n,i} \neq 0$.

Par sommabilité de $(S_i)_{i\in I}$, l'ensemble des $i \in I$ tels que $a_{n,i} \neq 0$ est fini. Comme $(I_\lambda)_{\lambda\in\Lambda}$ est une partition de I, il existe un sous-ensemble $\Lambda' \subset \Lambda$ fini tel que $i \in \bigcup_{\lambda\in\Lambda'} I_\lambda$. Pour tout indice $\lambda \notin \Lambda'$, on a donc $\sum_{i\in I_\lambda} a_{n,i} = 0$, d'où la sommabilité de la famille $\big(\sum_{i\in\lambda} S_i\big)_{\lambda\in\Lambda}$. On a alors

$$\sum_{\lambda\in\Lambda}\Big(\sum_{i\in I_\lambda} S_i\Big) = \sum_{n\geqslant 0}\Big(\sum_{\lambda\in\Lambda}\Big(\sum_{i\in I_\lambda} a_{n,i}\Big)\Big)X^n.$$

Mais, $(I_\lambda)_{\lambda\in\Lambda}$ est une partition de I, et ainsi

$$\sum_{\lambda\in\Lambda}\Big(\sum_{i\in I_\lambda} a_{n,i}\Big) = \sum_{i\in I} a_{n,i}.$$

Cela achève la démonstration. □

D'après l'exemple 2.2 (3), la définition suivante a un sens.

2.4. Définition. Soit $T = \sum_{n\geqslant 0} a_n X^n \in A[[X]]$. Pour tout $S \in A[[X]]$ tel que $v_X(S) \geqslant 1$, on note $T \circ S$ la série formelle

$$T \circ S = \sum_{k\geqslant 0} a_k S^k.$$

On dit que $T \circ S$ est la *composée* de T par S, ou que $T \circ S$ est obtenue par substitution de S à X dans T. On dit encore que $T \circ S$ est l'évaluation de T en S. On la note encore aussi parfois $T(S)$.

2.5. Exemples

(1) Si $T \in A[[X]]$ est une série formelle constante, alors $T(S) = T$ pour toute série formelle $S \in A[[X]]$ telle que $v_X(S) \geqslant 1$.

(2) Par définition, l'évaluation de $T \in A[[X]]$ en la série formelle X n'est rien d'autre que T.

(3) On peut toujours composer $T \in A[[X]]$ par la série formelle 0. La série composée $T(0)$ est alors simplement le terme constant de T.

En revanche, il n'est pas possible en général d'évaluer une série formelle en un élément $a \in A$, contrairement au cas des polynômes, simplement parce que la somme d'une famille infinie d'éléments de A n'a pas de sens.

La composition de séries formelles a de bonnes propriétés. C'est l'objet de la proposition de la page suivante.

2.6. Proposition. *Soit $S \in A[[X]]$ vérifiant $v_X(S) \geqslant 1$. Alors, l'application*

$$\begin{aligned} ev_S : A[[X]] &\longrightarrow A[[X]] \\ T &\longmapsto T \circ S \end{aligned}$$

est un morphisme d'anneaux.

En particulier, pour tout $a \in A$, et tout $T \in A[[X]]$, on a

$$(aT) \circ S = a(T \circ S).$$

Démonstration. Le seul point non évident à démontrer est l'égalité

$$(UV) \circ S = (U \circ S)(V \circ S), \quad \text{pour tous } U, V \in A[[X]].$$

Si $U = \sum_{n \geqslant 0} a_n X^n$ et $V = \sum_{n \geqslant 0} b_n X^n$, on a

$$(UV) \circ S = \sum_{n \geqslant 0} \Big(\sum_{m=0}^{n} a_m b_{n-m} \Big) S^n,$$

soit

$$(UV) \circ S = \sum_{n \geqslant 0} \Big(\sum_{m=0}^{n} a_m S^m b_{n-m} S^{n-m} \Big).$$

Pour tous p, $q \geqslant 0$, on a $v_X(a_p b_q S^{p+q}) \geqslant p+q$. Pour tout $n \geqslant 0$, l'ensemble des couples $(p, q) \in \mathbb{N}^2$ tels que $n \geqslant p + q$ est fini. On déduit aisément que la famille $(a_p b_q S^{p+q})_{(p,q) \in \mathbb{N}^2}$ est sommable. Considérons la partition

$$\mathbb{N}^2 = \bigcup_{n \geqslant 0} \{(p, q) \in \mathbb{N}^2 \mid p + q = n\};$$

la proposition 2.3 (3) donne alors

$$\sum_{(p,q) \in \mathbb{N}^2} a_p b_q S^{p+q} = \sum_{n \geqslant 0} \sum_{p+q=n} a_p b_q S^{p+q} = \sum_{n \geqslant 0} \Big(\sum_{m=0}^{n} a_m S^m b_{n-m} S^{n-m} \Big),$$

par commutativité de $A[[X]]$. Par suite, on a

$$\sum_{(p,q) \in \mathbb{N}^2} a_p b_q S^{p+q} = (UV) \circ S.$$

Par ailleurs, en considérant cette fois la partition

$$\mathbb{N}^2 = \bigcup_{q \geqslant 0} \{(p, q) \mid p \geqslant 0\},$$

on obtient

$$\sum_{(p,q) \in \mathbb{N}^2} a_p b_q S^{p+q} = \sum_{q \geqslant 0} \Big(\sum_{p \geqslant 0} a_p b_q S^{p+q} \Big),$$

en appliquant le point (3) de la proposition 2.3 (3).

D'après le point (2) de cette même proposition, appliqué à la famille $(a_pS_p)_{p\geqslant 0}$ et $T = b_qS^q$ donne alors

$$\sum_{(p,q)\in\mathbb{N}^2} a_pb_qS^{p+q} = \sum_{q\geqslant 0}\Big(\sum_{p\geqslant 0} a_pS^p\Big)b_qS^q.$$

En appliquant ce même point à $(b_qS^q)_{q\geqslant 0}$ et $T = \big(\sum_{p\geqslant 0} a_pS^p\big)$, on obtient

$$\sum_{(p,q)\in\mathbb{N}^2} a_pb_qS^{p+q} = \Big(\sum_{q\geqslant 0} b_qS^q\Big)\Big(\sum_{p\geqslant 0} a_pS^p\Big) = (V\circ S)(U\circ S).$$

En utilisant la commutativité de $A[[X]]$, on obtient l'égalité désirée.

Si maintenant $a\in A$ et $T\in A[[X]]$, on a

$$(aT)\circ S = (a\circ S)(T\circ S) = a(T\circ S).$$

Cela achève la démonstration. □

2.7. Exemples

(1) Ce résultat permet par exemple de donner une expression de l'inverse d'une série formelle $F\in A[[X]]^\times$. D'après le lemme 1.4, le terme constant a_0 de F est inversible. Écrivons $T = a_0 + S$, avec $v_X(S)\geqslant 1$. On a alors $T = a_0(1 + a_0^{-1}S)$. Remarquons maintenant que

$$v_X(-a_0^{-1}S) = v_X(S)\geqslant 1.$$

Une substitution dans la relation

$$(1-X)\sum_{n\geqslant 0} X^n = 1$$

donne

$$(1+a_0^{-1}S)\Big(\sum_{n\geqslant 0} a_0^{-n}(-1)^nS^n\Big) = 1.$$

On en déduit alors les égalités

$$T^{-1} = \sum_{n\geqslant 0} a_0^{-n-1}(-1)^nS^n = \sum_{n\geqslant 0} a_0^{-n-1}(-1)^n(T-a_0)^n.$$

(2) L'évaluation en S étant un morphisme d'anneaux, pour toute série formelle $T\in A[[X]]^\times$ et toute $S\in A[[X]]$ telle que $v_X(S)\geqslant 1$, la série formelle $T(S)$ est inversible, et l'on a

$$\big(T(S)\big)^{-1} = T^{-1}(S).$$

Par exemple, les séries formelles $-X, X^2, -X^2\in A[[X]]$ sont d'ordre au moins 1.

Par substitution, l'égalité

$$(1-X)^{-1} = \sum_{n\geqslant 0} X^n$$

donne alors les égalités

$$(1+X)^{-1} = \sum_{n\geqslant 0}(-1)^n X^n,$$

$$(1-X^2)^{-1} = \sum_{n\geqslant 0} X^{2n}$$

$$(1+X^2)^{-1} = \sum_{n\geqslant 0}(-1)^n X^{2n}.$$

Notons que l'on peut aussi retrouver ces identités en utilisant l'exemple précédent.

2.8. Lemme. *Soient $S, T \in A[[X]]$. On suppose que $v_X(S) \geqslant 1$. Alors,*

$$v_X(T\circ S) \geqslant v_X(T)v_X(S),$$

avec égalité si A est intègre.

Démonstration. Si $k = v_X(T)$, on écrit $T = \sum_{n\geqslant k} b_n X^n$. On a alors

$$T\circ S = \sum_{n\geqslant k} b_n S^n.$$

Or, d'après le lemme 1.2 (4), pour tout $n \geqslant k$, on a

$$v_X(b_n S^n) \geqslant v_X(b_n) + nv_X(S) \geqslant nv_X(S).$$

En particulier, pour tout $n \geqslant k$, on a

$$v_X(b_n S^n) \geqslant v_X(S)v_X(T).$$

Ainsi, pour tout $n \geqslant k$, les coefficients de $b_n S^n$ d'indice $< v_X(S)v_X(T)$ sont nuls. Il en est donc de même pour $T\circ S$.

Supposons A intègre. Pour tout $n \geqslant k+1$, on a

$$v_X(b_n S^n) \geqslant nv_X(S) > kv_X(S),$$

car $v_X(S) \geqslant 1$. On en déduit alors que $v_X\big(\sum_{n\geqslant k+1} b_n S^n\big) > kv_X(S)$. De plus, on a $v_X(b_k S^k) = kv_X(S)$ d'après le lemme 1.2 (4). Mais alors, par le point (3) de ce même lemme, on a

$$v_X(T\circ S) = v_X\Big(b_k S^k + \sum_{n\geqslant k+1} b_n S^n\Big) = kv_X(S) = v_X(S)v_X(T).$$

Cela achève la démonstration. □

Remarquons que $XA[[X]]$ est l'ensemble des séries formelles $S \in A[[X]]$ vérifiant $v_X(S) \geqslant 1$.

Le lemme 2.8 montre en particulier que si $S, T \in XA[[X]]$, il en est de même de $T \circ S$. L'énoncé du corollaire suivant a donc du sens.

2.9. Corollaire. *Pour tous $U, T, S \in A[[X]]$ avec $S, T \in XA[[X]]$, on a*

$$U \circ (T \circ S) = (U \circ T) \circ S.$$

En particulier, la composition des séries formelles induit une loi interne associative sur $XA[[X]]$, de neutre X.

Démonstration. Écrivons $U = \sum_{n \geqslant 0} c_n X^n$. On suppose d'abord que U est un polynôme. Dans ce cas, la somme est finie. D'après la proposition 2.6, on a alors

$$\begin{aligned} U \circ (T \circ S) &= \sum_{n \geqslant 0} c_n (T \circ S)^n \\ &= \sum_{n \geqslant 0} c_n (T^n \circ S) \\ &= \sum_{n \geqslant 0} (c_n T^n) \circ S \\ &= \Big(\sum_{n \geqslant 0} c_n T^n \Big) \circ S \\ &= (U \circ T) \circ S. \end{aligned}$$

Passons maintenant au cas général. Pour tout $n \geqslant 0$, on écrit $U = U_n + R_n$, où $U_n \in A[X]$ est un polynôme de degré $\leqslant n-1$ et $v_X(R_n) \geqslant n$. En utilisant le fait que $ev_{T \circ S}$ et $ev_T \circ ev_S$ sont des morphismes d'anneaux, pour tout entier $n \geqslant 0$, on a alors

$$U \circ (T \circ S) - (U \circ T) \circ S = U_n \circ (T \circ S) - (U_n \circ T) \circ S + W_n = W_n,$$

avec $W_n = R_n \circ (T \circ S) - (R_n \circ T) \circ S$, la dernière égalité provenant du point précédent. Or, le lemme 2.8 montre que

$$v_X\big(R_n \circ (T \circ S)\big) \geqslant v_X(R_n) v_X(T \circ S) \geqslant v_X(R_n) v_X(T) v_X(S) \geqslant n.$$

De même, on a

$$v_X\big((R_n \circ T) \circ S\big) \geqslant v_X(R_n \circ T) v_X(S) \geqslant v_X(R_n) v_X(T) v_X(S) \geqslant n.$$

Par conséquent, on a $v_X(W_n) \geqslant n$. Autrement dit, on a

$$v_X\big(U \circ (T \circ S) - (U \circ T) \circ S\big) \geqslant n, \quad \text{pour tout } n \geqslant 0.$$

Cela signifie simplement que $U \circ (T \circ S) - (U \circ T) \circ S = 0$.

Le reste de la proposition est alors clair. □

On s'intéresse maintenant à la généralisation des morphismes d'évaluation d'un polynôme au cas des séries formelles. Comme on l'a déjà remarqué, ce n'est pas toujours possible. En revanche, tout se passe bien lorsque l'on évalue en un élément nilpotent.

2.10. Proposition. *Soit B un anneau commutatif, et soit A un sous-anneau de B. Pour tout $b \in \mathrm{Nil}(B)$, l'application*

$$\begin{aligned} ev_b : \quad A[[X]] &\longrightarrow B \\ S = \sum_{n\geqslant 0} a_n X^n &\longmapsto S(b) = \sum_{n\geqslant 0} a_n b^n \end{aligned}$$

est un morphisme d'anneaux bien défini. De plus, on a

$$\mathrm{Im}(ev_b) = A[b] = \{P(b) \mid P \in A[X]\}.$$

Enfin, pour tout $T \in XA[[X]]$, $T(b) \in \mathrm{Nil}(B)$ et l'on a

$$(S \circ T)(b) = S\big(T(b)\big), \qquad \textit{pour tout } S \in A[[X]].$$

Démonstration. Puisque $b \in \mathrm{Nil}(B)$, la somme $\sum_{n\geqslant 0} a_n b^n$ est finie. Par conséquent, ev_b est bien définie et l'égalité $\mathrm{Im}(ev_b) = A[b]$ est alors claire. Le fait que ev_b soit un morphisme se montre comme la proposition XV-6.16.

Soit $T \in XK[[X]]$. Écrivons $T = XU, U \in A[[X]]$. D'après le premier point, pour tout $n \geqslant 0$, on a

$$T(b)^n = T^n(b) = (X^n U^n)(b) = b^n U^n(b).$$

En particulier, si $r \geqslant 1$ vérifie $b^r = 0$, on obtient $T(b)^r = 0$, et l'on a bien ainsi $T(b) \in \mathrm{Nil}(B)$.

Si maintenant $S = \sum_{n\geqslant 0} a_n X^n$, on a alors

$$S\big(T(b)\big) = \sum_{n\geqslant 0} a_n T(b)^n = \sum_{n=0}^{r-1} a_n T(b)^n = \Big(\sum_{n=0}^{r-1} a_n T^n\Big)(b).$$

Par ailleurs, pour tout $n \geqslant r$, on a

$$v_X(T^n) \geqslant n v_X(T) \geqslant n \geqslant r.$$

Il existe donc $V \in K[[X]]$ tel que

$$S \circ T = \sum_{n\geqslant 0} a_n T^n = \sum_{n=0}^{r-1} a_n T^n + X^r V.$$

Toujours d'après le premier point, on a alors

$$(S \circ T)(b) = \Big(\sum_{n=0}^{r-1} a_n T^n\Big)(b) + b^r V(b) = \Big(\sum_{n=0}^{r-1} a_n T^n\Big)(b),$$

d'où l'égalité $(S \circ T)(b) = S\big(T(b)\big)$. Cela achève la démonstration. □

On va maintenant étudier les conditions d'existence d'un inverse pour la composition. Remarquons que, pour pouvoir composer deux séries formelles S et T dans les deux sens, il faut que $S, T \in XA[[X]]$.

2.11. Théorème. *Soit $S \in XA[[X]]$. Alors, S admet un symétrique pour la composition dans $XA[[X]]$ si, et seulement si, son coefficient en X est inversible dans A. Dans ce cas, S et son symétrique sont d'ordre 1, et leurs termes en X sont inversibles dans A.*

En particulier, l'ensemble des séries formelles d'ordre 1 dont le terme en X est inversible est un groupe pour la composition, de neutre X.

Démonstration. Soit $S = \sum_{n\geqslant 1} a_n X^n \in XA[[X]]$. Alors, on a les propriétés suivantes :

(i) pour tout $k \geqslant 1$, le coefficient en X^k de S^k est a_1^k ;
(ii) pour tout $k \geqslant 2$, S^k n'a pas de terme en X ;
(iii) pour tout $T = \sum_{n\geqslant 1} b_n X^n \in XA[[X]]$, le coefficient en X de $T \circ S$ est égal à $b_1 a_1$.

En effet, on a $S = X(a_1 + XS_1)$, avec $S_1 \in A[[X]]$. En appliquant la formule du binôme, on obtient ainsi

$$S^k = X^k(a_1^k + XS_2),$$

avec $S_2 \in A[[X]]$. Les points (i) et (ii) en découlent immédiatement.

Si $T = \sum_{n\geqslant 1} b_n X^n \in XA[[X]]$, alors d'après (ii), S^k n'a donc pas de terme en X pour tout $k \geqslant 2$, et il est donc de même de $b_k S^k$. Ainsi, seul $b_1 S$ contribue au terme en X. Le coefficient en X de $T \circ S$ est donc celui de $b_1 S$, c'est-à-dire $b_1 a_1$, d'où (iii).

Supposons maintenant que $T = \sum_{n\geqslant 1} b_n X^n \in XA[[X]]$ vérifie $T \circ S = X$. En utilisant (iii) et en comparant les termes en X, on obtient $b_1 a_1 = 1$, et par conséquent $a_1 \in A^\times$ (puisque A est commutatif). Comme le coefficient en X de T est b_1, on a $v_X(T) = 1$ et le coefficient en X de T est aussi inversible.

Nous allons maintenant montrer que pour tout $S \in XA[[X]]$ dont le coefficient en X est inversible dans A, il existe $T = \sum_{n\geqslant 1} b_n X^n \in XA[[X]]$ tel que l'on ait $T \circ S = X$.

Pour tout $S \in XA[[X]]$, et pour tout $k \geqslant 1$, on écrit $S^k = \sum_{n\geqslant 1} a_n^{(k)} X^n$. Puisque $v_X(S^k) \geqslant k v_X(S) \geqslant k$, on a $a_n^{(k)} = 0$ dès que $k > n$. On a alors

$$T \circ S = \sum_{n\geqslant 1} \Big(\sum_{k\geqslant 1} b_k a_n^{(k)} \Big) X^n = \sum_{n\geqslant 1} \Big(\sum_{k=1}^{n} b_k a_n^{(k)} \Big) X^n.$$

On doit donc vérifier qu'il existe une suite $(b_n)_{n\geqslant 1}$ d'éléments de A telle que

$$b_1 a_1 = 1 \quad \text{et} \quad \sum_{k=1}^{n} b_k a_n^{(k)} = 0, \qquad \text{pour tout } n \geqslant 2.$$

Par (i), on a $a_n^{(n)} = a_1^n$ pour tout $n \geqslant 1$. Les relations précédentes se récrivent donc

$$b_1 a_1 = 1 \quad \text{et} \quad \sum_{k=1}^{n-1} b_k a_n^{(k)} + b_n a_1^n = 0, \qquad \text{pour tout } n \geqslant 2,$$

soit encore

$$b_1 = a_1^{-1} \quad \text{et} \quad b_n = -a_1^{-n} \Big(\sum_{k=1}^{n-1} b_k a_n^{(k)} \Big).$$

Une récurrence immédiate montre alors que la suite $(b_n)_{n\geqslant 1}$ recherchée existe, et il existe donc bien $T \in XA[[X]]$ telle que $T \circ S = X$.

Montrons maintenant la proposition.

Si $S \in XA[[X]]$ admet un symétrique $T \in XA[[X]]$ pour la composition, alors en particulier $T \circ S = X$, et par un point précédent, le terme en X de S est inversible dans A.

Inversement, si le terme en X de S est inversible dans A, on vient de voir qu'il existe $T \in XA[[X]]$ telle que $T \circ S = X$. De plus, on a remarqué que dans ce cas, le coefficient en X de T est inversible. Il existe alors une série formelle $U \in XA[[X]]$ tel que $U \circ T = X$. Mais alors, on a

$$(U \circ T) \circ S = X \circ S = S \quad \text{et} \quad U \circ (T \circ S) = U \circ X = U.$$

Par le corollaire 2.9, on obtient $S = U$, d'où $S \circ T = X$. Ainsi, S est inversible dans $XA[[X]]$ d'inverse T.

Montrons enfin la dernière partie. Soit $\mathscr{S}$ l'ensemble des séries formelles d'ordre 1 et donc le coefficient de X est inversible dans A. Si $S, T \in \mathscr{S}$, le point (iii) montre qu'il en est de même de $T \circ S$. Ainsi, la composition induit une loi interne sur $\mathscr{S}$. De plus, la loi de composition est associative et de neutre X, d'après le corollaire 2.9. Enfin, les considérations précédentes montrent que tout élément de $\mathscr{S}$ admet un symétrique dans $\mathscr{S}$. Cela achève la démonstration. □

Puisque tout élément non nul d'un corps est inversible, on obtient le résultat suivant.

2.12. Corollaire. *Soit K un corps. Alors, l'ensemble des séries formelles de $K[[X]]$ d'ordre 1 est un groupe pour la composition, de neutre X.*

2.13. Définition. Soit $S \in A[[X]]$ une série formelle d'ordre 1, dont le coefficient en X inversible. Le symétrique de S pour la composition est noté $S^{\langle -1\rangle}$, et appelé la série *réverse* de S.

Nous allons maintenant nous intéresser à une autre opération sur les séries formelles, à savoir la dérivation.

2.14. Définition. Soit $S = \sum_{n\geqslant 0} a_n X^n \in A[[X]]$. La série formelle

$$S' = \sum_{n\geqslant 1} n a_n X^{n-1} = \sum_{n\geqslant 0} (n+1) a_{n+1} X^n$$

est appelée la *dérivée (formelle)* de S.

La dérivation des séries formelles a des propriétés naturelles.

2.15. Proposition. *Soient $S, T \in A[[X]]$ et soit $a \in A$. Alors :*

(1) $(aS)' = aS'$;

(2) $(S+T)' = S' + T'$;

(3) $(ST)' = S'T + ST'$;

(4) *pour tout $k \geqslant 1$, on a $(S^k)' = kS^{k-1}S'$;*

(5) *si $v_X(S) \geqslant 1$, on a $(T \circ S)' = (T' \circ S)S'$.*
Si de plus le coefficient en X de S est inversible, alors $S' \circ S^{\langle -1\rangle}$ est inversible dans $A[[X]]$, et l'on a

$$(S^{\langle -1\rangle})' = (S' \circ S^{\langle -1\rangle})^{-1}.$$

Démonstration. Les propriétés (1) – (3) découlent de calculs immédiats laissés au lecteur, et (4) s'obtient par récurrence en utilisant (3).

Montrons (5). Écrivons $T = \sum_{n\geqslant 0} b_n X^n$. On a donc $T \circ S = b_0 + \sum_{k\geqslant 1} b_k S^k$. On suppose tout d'abord que T est un polynôme. Dans ce cas, cette dernière somme est finie et les points précédents montrent que l'on a

$$(T \circ S)' = \Big(\sum_{k\geqslant 1} k b_k S^{k-1}\Big) S' = (T' \circ S) S'.$$

Revenons au cas général. Pour tout entier $n \geqslant 0$, on écrit $T = U + X^n R_n$, avec $R_n \in A[[X]]$ et où $U \in A[X]$ est de degré $\leqslant n-1$.

On a alors $$T \circ S = U \circ S + S^n(R_n \circ S),$$

d'après la proposition 2.6. D'après ce qui précède et les points (1) – (4), on a alors

$$(T \circ S)' = (U' \circ S)S' + nS^{n-1}S'(R_n \circ S) + S^n(R_n \circ S)'.$$

Par ailleurs, on a $T' = U' + nX^{n-1}R_n + X^nR_n'$, et ainsi

$$(T' \circ S)S' = (U' \circ S)S' + nS^{n-1}(R_n \circ S)S' + S^n(R_n' \circ S)S',$$

d'où $$(T \circ S)' - (T' \circ S)S' = S^n\big((R_n \circ S)' - (R_n' \circ S)S'\big).$$

Mais alors, pour tout $n \geqslant 0$, on a

$$v_X\big((T \circ S)' - (T' \circ S)S'\big) \geqslant v_X(S^n) \geqslant nv_X(S) \geqslant n.$$

Cela implique que $(T \circ S)' - (T' \circ S)S' = 0$ d'où la première partie.

Si maintenant le coefficient de S est inversible, il admet une série réverse $S^{\langle -1\rangle}$. Par définition, on a $S \circ S^{\langle -1\rangle} = X$. En utilisant ce qui précède on obtient $$(S' \circ S^{\langle -1\rangle})(S^{\langle -1\rangle})' = 1,$$

d'où le résultat voulu. □

Le lemme suivant sera très utile dans la suite.

2.16. Lemme. *Soit K un corps de caractéristique nulle. Si $S \in K[[X]]$ vérifie $S' = 0$ et $S(0) = 0$, alors $S = 0$. En particulier :*

(1) *deux séries formelles sont égales si, et seulement si, elles ont même dérivée et même terme constant ;*

(2) *pour tout $T \in K[[X]]$, il existe une unique série formelle $S \in XK[[X]]$ telle que $S' = T$.*

Démonstration. Soit $S = \sum_{n \geqslant 0} a_n X^n$ telle que $S' = 0$ et $S(0) = 0$. On a alors

$$a_0 = 0 \quad \text{et} \quad na_n = 0, \quad \text{pour tout } n \geqslant 1.$$

Comme K est de caractéristique nulle, on en déduit immédiatement que l'on a $a_n = 0$ pour tout $n \geqslant 0$, c'est-à-dire $S = 0$.

Le point (1) en découle alors en utilisant les propriétés de linéarité de la dérivation formelle. Enfin, si $T = \sum_{n \geqslant 0} b_n X^n$, la série formelle

$$S = \sum_{n \geqslant 0} \frac{b_n}{n+1} X^{n+1}$$

appartient à $XK[[X]]$ et vérifie $S' = T$.

L'unicité d'une série formelle vérifiant les conditions du point (2) est une conséquence immédiate du point (1). Cela achève la démonstration. □

Nous allons maintenant démontrer la formule de réversion de Lagrange, qui permet de calculer $S^{\langle -1\rangle}$.

Avant d'en donner l'énoncé, faisons tout d'abord quelques remarques préliminaires. Soit K un corps de caractéristique nulle, et soit $S \in K[[X]]$ d'ordre 1. On peut ainsi écrire $S = XT$. On voit aisément que $v_X(T) = 0$, et donc que T est inversible dans $K[[X]]$. De plus, le corollaire 2.12 montre que S admet une série reverse $S^{\langle -1\rangle}$, qui est aussi d'ordre 1. Pour tout entier $n \geqslant 1$, $(S^{\langle -1\rangle})^n$ est alors d'ordre n, et n'a donc pas de termes en X^k, pour tout $k \in [\![0, n-1]\!]$.

Notation. Si $S \in A[[X]]$, pour tout $n \geqslant 0$, on note $c_n(S)$ le coefficient de S en X^n. Autrement dit, si $S = \sum_{n \geqslant 0} a_n X^n$, on a $c_n(S) = a_n$.

On a alors le résultat suivant.

2.17. Théorème (Schur). *Soit K un corps de caractéristique nulle.*

Soit $S \in K[[X]]$ une série formelle d'ordre 1, et soit $T \in K[[X]]$ l'unique série formelle telle que $S = XT$. Alors, pour tout $n \geqslant 1$, on a

$$(S^{\langle -1\rangle})^n = \sum_{k \geqslant n} \frac{n}{k} c_{k-n}(T^{-k}) X^k.$$

Démonstration. Pour tout $F \in K[[X]]$, et tout $k \geqslant 0$, on note $c_k(F) \in K$ le coefficient en X^k de F.

Pour tous $k \geqslant n \geqslant 1$, posons $c_k(S^n) = a_{k,n}$ et $c_k\big((S^{\langle -1\rangle})^n\big) = b_{k,n}$, si bien que

$$S^n = \sum_{k \geqslant n} a_{k,n} X^k, \quad (S^{\langle -1\rangle})^n = \sum_{k \geqslant n} b_{k,n} X^k.$$

Fixons $n \geqslant 1$. D'après la proposition 2.6, on a

$$S^n \circ S^{\langle -1\rangle} = (S \circ S^{\langle -1\rangle})^n = X^n.$$

On a donc

$$X^n = \sum_{m \geqslant n} a_{m,n} (S^{\langle -1\rangle})^m = \sum_{m \geqslant n} a_{m,n} \Big(\sum_{k \geqslant m} b_{k,m} X^k \Big).$$

Or, on a l'égalité

$$\{(m,k) \mid m \geqslant n \geqslant 1,\ k \geqslant m\} = \{(m,k) \mid k \geqslant n \geqslant 1,\ m \in [\![n,k]\!]\},$$

et l'on obtient ainsi

$$X^n = \sum_{k\geqslant n} \Big(\sum_{m=n}^{k} a_{m,n} b_{k,m} \Big) X^k.$$

Par identification, on a

$$a_{n,n} b_{n,n} = 1 \quad \text{et} \quad \sum_{m=n}^{k} a_{m,n} b_{k,m} = 0, \quad \text{pour tout } k > n. \qquad (R_n)$$

Comme déjà noté dans la démonstration du théorème 2.11, le coefficient en X^n de S^n est $a_{1,1}^n$, où $a_{1,1} = c_1(S)$. On a donc $a_{n,n} = a_{1,1}^n \in K^\times$.

Une récurrence immédiate montre alors que $(b_{k,m})_{k\geqslant m\geqslant 1}$ est l'unique suite double d'éléments de K vérifiant les relations (R_n) pour tout $n \geqslant 1$.

Pour tout $k \geqslant m \geqslant 1$, posons $b'_{k,m} = \dfrac{m}{k} c_{k-m}(T^{-k})$. Nous allons montrer que la suite double $(b'_{k,m})_{k\geqslant m\geqslant 1}$ vérifie les relations (R_n) pour tout $n \geqslant 1$, ce qui permettra de conclure d'après la remarque précédente.

Pour tout $n \geqslant 1$, $b'_{n,n}$ est le coefficient constant de la série formelle T^{-n}. Or, on a $S^n = X^n T^n$, d'où $T^{-n} S^n = X^n$. Comme S^n est d'ordre n, le coefficient de X^n du produit $S^n T^{-n}$ est donc $a_{n,n} b'_{n,n}$. Ainsi, $a_{n,n} b'_{n,n} = 1$.

Si $k > n$, posons $d_{k,n} = \sum_{m=n}^{k} a_{m,n} b'_{k,m}$. On doit vérifier que $d_{k,n} = 0$. On commence par remarquer que

$$\begin{aligned} X(S^n)' T^{-k} &= \Big(\sum_{i\geqslant n} i a_{i,n} X^i \Big) \Big(\sum_{j\geqslant 0} c_j(T^{-k}) X^j \Big) \\ &= \sum_{\ell\geqslant n} \Big(\sum_{m=n}^{\ell} m a_{m,n} c_{\ell-m}(T^{-k}) \Big) X^\ell. \end{aligned}$$

En particulier, on a

$$c_k\big(X(S^n)' T^{-k}\big) = \sum_{m=n}^{k} m a_{m,n} c_{k-m}(T^{-k}) = \sum_{m=n}^{k} k a_{m,n} b'_{k,m}.$$

Ainsi, $d_{k,n} = \dfrac{1}{k} c_k\big(X(S^n)' T^{-k}\big)$. Or, on a

$$X(S^n)' T^{-k} = nXT^{-k} S^{n-1} S' = nXT^{-k} X^{n-1} T^{n-1} (XT' + T),$$

soit

$$X(S^n)' T^{-k} = nX^n T^{n-k-1} (XT' + T).$$

Par conséquent,

$$\begin{aligned} d_{k,n} &= \frac{n}{k} c_k\big(X^n T^{n-k-1} (XT' + T)\big) \\ &= \frac{n}{k} c_{k-n}\big(T^{n-k-1} (XT' + T)\big) \\ &= \frac{n}{k} c_{k-n}(T^{n-k} + XT' T^{n-k-1}). \end{aligned}$$

Si l'on pose $U = T^{n-k}$, on a par conséquent

$$d_{k,n} = \frac{n}{k} c_{k-n}\big(U - \frac{1}{k-n} XU'\big).$$

Or, un simple calcul montre que pour toute série formelle $F \in K[[X]]$ et tout $k > n$, le coefficient en X^{k-n} de $F - \frac{1}{k-n} XF'$ est nul. Cela achève la démonstration. □

On a alors immédiatement le corollaire suivant.

2.18. Corollaire (Formule de réversion de Lagrange). *Soit K un corps de caractéristique nulle.*

Soit $S \in K[[X]]$ tel que $v_X(S) = 1$, soit $T \in K[[X]]$ telle que $S = XT$. Alors, on a

$$S^{\langle -1\rangle} = \sum_{k \geqslant 1} \frac{1}{k} c_{k-1}(T^{-k}) X^k.$$

Nous verrons des applications de la formule de réversion de Lagrange un peu plus loin. Nous terminons ce paragraphe en généralisant cette formule.

2.19. Théorème. *Soit K un corps de caractéristique nulle.*

Soit S une série formelle d'ordre 1 à coefficients dans K, et soit $T \in K[[X]]$ tel que $S = XT$. Pour toute $W \in K[[X]]$ tel que $v_X(W) \geqslant 1$, on a

$$W \circ S^{\langle -1\rangle} = \sum_{k \geqslant 1} \frac{1}{k} c_{k-1}(W'T^{-k}) X^k.$$

Démonstration. Commençons par le cas $W = X^n$, où $n \geqslant 1$. Dans ce cas, on a $W' = nX^{n-1}$. On remarque alors que $W'T^{-k}$ est d'ordre $n-1$. En particulier, $c_{k-1}(W'T^{-k}) = 0$ pour tout $k \in [\![1, n-1]\!]$. D'autre part, pour tout $k \geqslant n$, on a

$$c_{k-1}(W'T^{-k}) = n c_{k-n}(T^{-k}).$$

Comme $W \circ S^{\langle -1\rangle} = (S^{\langle -1\rangle})^n$, on doit donc établir l'égalité

$$(S^{\langle -1\rangle})^n = \sum_{k \geqslant n} \frac{n}{k} c_{k-n}(T^{-k}) X^k,$$

ce qui est exactement l'énoncé du théorème de Schur.

En utilisant les propriétés de l'évaluation et de la dérivation, il s'ensuit que le théorème est également vrai pour tout polynôme $W \in K[X]$ sans terme constant. Si $W \in K[[X]]$ est d'ordre $v_X(W) \geqslant 1$, pour tout entier $n \geqslant 1$, on écrit $W = W_n + R_n$, où W_n est un polynôme de degré $< n$ sans terme constant et où $R_n \in K[[X]]$ vérifie $v_X(R_n) \geqslant n$.

Pour chaque $V \in K[[X]]$ telle que $v_X(V) \geqslant 1$, posons

$$H_V = \sum_{k \geqslant 1} \frac{1}{k} c_{k-1}(V'T^{-k})X^k.$$

On a ainsi

$$W \circ S^{\langle -1 \rangle} = W_n \circ S^{\langle -1 \rangle} + R_n \circ S^{\langle -1 \rangle}.$$

D'après le point précédent, on a $W_n \circ S^{\langle -1 \rangle} = H_{W_n} = H_W - H_{R_n}$. Par conséquent,

$$W \circ S^{\langle -1 \rangle} = H_W - H_{R_n} + R_n \circ S^{\langle -1 \rangle},$$

soit

$$W \circ S^{\langle -1 \rangle} - H_W = -H_{R_n} + R_n \circ S^{\langle -1 \rangle}.$$

Or, on a

$$v_X(R_n \circ S^{\langle -1 \rangle}) \geqslant v_X(R_n) v_X(S^{\langle -1 \rangle}) \geqslant n$$

d'après le lemme 2.8. Comme $v_X(R_n) \geqslant n$, on a $v_X(R_n') \geqslant n-1$. Ainsi, on en déduit que $v_X(H_{R_n}) \geqslant n-1$. D'après le lemme 1.2, on obtient ainsi

$$v_X(W \circ S^{\langle -1 \rangle} - H_W) \geqslant n-1, \quad \text{pour tout } n \geqslant 1,$$

d'où $W \circ S^{\langle -1 \rangle} - H_W = 0$, ce qu'il fallait démontrer. □

2.20. Corollaire. *Soit K un corps de caractéristique nulle.*

Soit $S \in K[[X]]$ d'ordre 1, et soit $T \in K[[X]]$ tel que $S = XT$. Pour toute série formelle $U \in K[[X]]$, on a

$$(U \circ S^{\langle -1 \rangle})(S^{\langle -1 \rangle})' = \sum_{k \geqslant 0} c_k(UT^{-k-1})X^k.$$

Démonstration. Soit $U \in K[[X]]$. Compte tenu du lemme 2.16 (2), on peut écrire $U = W'$, où $W \in K[[X]]$ est une série formelle d'ordre $v_X(W) \geqslant 1$. D'après le théorème 2.19, on a

$$W \circ S^{\langle -1 \rangle} = \sum_{k \geqslant 1} \frac{1}{k} c_{k-1}(W'T^{-k})X^k.$$

Par dérivation, on obtient

$$(W' \circ S^{\langle -1 \rangle})(S^{\langle -1 \rangle})' = \sum_{k \geqslant 1} c_{k-1}(W'T^{-k})X^{k-1} = \sum_{k \geqslant 0} c_k(W'T^{-k-1})X^k,$$

d'où le résultat, puisque $U = W'$. □

3. Séries formelles usuelles

Nous allons ici donner quelques exemples de séries formelles usuelles [1]. Avant de le faire, nous allons démontrer un analogue d'un théorème de Cauchy pour les équations différentielles linéaires à coefficients dans l'anneau $K[[X]]$.

1. Enfin !!

Notations. Pour tout $n \geqslant 0$, et tout $S \in K[[X]]$, on note $S^{(n)} \in K[[X]]$ la série formelle obtenue par n dérivations successives. On a donc $S^{(0)} = S$.

De plus, on note aussi $S(0) \in K$ le terme constant de S.

Dans tout ce paragraphe, K désignera un corps de caractéristique nulle. On a alors le résultat suivant.

3.1. Proposition. *Soit $d \geqslant 1$ et soient $A_0, \ldots, A_d \in K[[X]]$. On suppose que $A_d \in K[[X]]^\times$. Alors, pour tous $y_0, \ldots, y_{d-1} \in K$, il existe une unique série formelle $S \in K[[X]]$ telle que*

$$A_d S^{(d)} + \cdots + A_1 S' + A_0 S = 0, \quad \text{et} \quad S^{(k)}(0) = y_k, \quad \text{pour tout } k \in [\![0, d-1]\!].$$

Démonstration. Écrivons $S = \sum_{n \geqslant 0} a_n X^n$. On a alors

$$S^{(\ell)} = \sum_{n \geqslant 0} (n+1)\cdots(n+\ell) a_{n+\ell} X^n = \sum_{n \geqslant 0} \frac{(n+\ell)!}{n!} a_{n+\ell} X^n, \quad \text{pour tout } \ell \geqslant 0.$$

En particulier, $S^{(\ell)}(0) = \ell!\, a_\ell$ pour tout $\ell \geqslant 0$.

Puisque A_d est inversible, l'équation se récrit

$$S^{(d)} = B_{d-1} S^{(d-1)} + \cdots + B_0 S, \quad \text{et} \quad S^{(k)}(0) = y_k, \quad \text{pour tout } k \in [\![0, d-1]\!],$$

avec $B_k = -A_d^{-1} A_k$ pour tout $k \in [\![0, d-1]\!]$.

Soit $k \in [\![0, d-1]\!]$. Si $B_k = \sum_{n \geqslant 0} b_{k,n} X^n$, on a alors

$$B_k S^{(k)} = \sum_{n \geqslant 0} \Big(\sum_{m=0}^{n} \frac{(m+k)!}{m!} a_{m+k} b_{k,n-m} \Big) X^n.$$

L'équation à laquelle nous nous intéressons se retraduit alors par

$$\begin{cases} \dfrac{(n+d)!}{n!} a_{n+d} &= \displaystyle\sum_{k=0}^{d-1} \Big(\sum_{m=0}^{n} \frac{(m+k)!}{m!} a_{m+k} b_{k,n-m} \Big), \quad \text{pour tout } n \geqslant 0 \\ k!\, a_k &= y_k, \quad \text{pour tout } k \in [\![0, d-1]\!]. \end{cases}$$

Le corps K étant de caractéristique nulle, cela se récrit enfin

$$\begin{cases} a_{n+d} &= \dfrac{n!}{(n+d)!} \displaystyle\sum_{k=0}^{d-1} \Big(\sum_{m=0}^{n} \frac{(m+k)!}{m!} a_{m+k} b_{k,n-m} \Big), \quad \text{pour tout } n \geqslant 0 \\ a_k &= \dfrac{y_k}{k!}, \quad \text{pour tout } k \in [\![0, d-1]\!]. \end{cases}$$

Une récurrence immédiate montre qu'il existe une unique suite $(a_n)_{n \geqslant 0}$ d'éléments de K vérifiant cette relation de récurrence linéaire, d'où le résultat souhaité. □

La première série formelle classique que nous allons présenter est la *série exponentielle*, à savoir

$$e^X = \sum_{n\geqslant 0} \frac{X^n}{n!}.$$

C'est l'unique série formelle $F \in K[[X]]$ solution de l'équation différentielle

$$F' - F = 0, \quad F(0) = 1.$$

Remarquons qu'ici, on a identifié un élément de $\mathbb{Q}$ à son image dans K. Nous ferons cet abus de notation bien pratique également dans la suite.

On définit maintenant la série logarithmique

$$\ln(1+X) = \sum_{n\geqslant 1} \frac{(-1)^{n-1}}{n} X^n.$$

On remarque que

$$\big(\ln(1+X)\big)' = \sum_{n\geqslant 1} (-1)^{n-1} X^{n-1} = \sum_{n\geqslant 0} (-1)^n X^n = (1+X)^{-1}.$$

Ainsi, $\ln(1+X)$ est l'unique série formelle telle que

$$F' = (1+X)^{-1}, \quad F(0) = 0.$$

Par composition, on peut donc définir e^S et pour tout $S \in XK[[X]]$, ainsi que $\ln(T)$ pour tout $T \in 1 + XK[[X]] = \{1 + S \mid S \in XK[[X]]\}$.

L'exponentielle et le logarithme ont de bonnes propriétés vis-à-vis de l'addition.

3.2. Proposition. *Les ensembles $XK[[X]]$ et $1 + XK[[X]]$ sont des sous-groupes de $K[[X]]$ et $K[[X]]^\times$, respectivement.*

De plus, l'application

$$\begin{array}{rccc} \exp : & XK[[X]] & \longrightarrow & 1 + XK[[X]] \\ & S & \longmapsto & e^S \end{array}$$

est un isomorphisme de groupes.

L'isomorphisme réciproque s'écrit

$$\begin{array}{rccc} \ln : & 1 + XK[[X]] & \longrightarrow & XK[[X]] \\ & T & \longmapsto & \ln(T). \end{array}$$

Démonstration. Le premier point découle de simples calculs, et est laissé au lecteur. Remarquons maintenant que pour tout $S \in XK[[X]]$, on a

$$v_X(S^n) = n v_X(S) \geqslant 1, \quad \text{pour tout } n \geqslant 1.$$

En particulier, le terme constant de e^S est 1, et ainsi $e^S \in 1 + XK[[X]]$. Cela implique au passage que e^S est inversible.

Soient maintenant S_1, $S_2 \in XK[[X]]$. Posons $U = e^{S+T}$ et $V = e^S e^T$. Puisque $(e^X)' = e^X$, on a

$$U' = (S+T)'e^{S+T} = (S'+T')U,$$

ainsi que

$$V' = (S'e^S)e^T + e^S(T'e^T) = (S'+T')e^S e^T = (S'+T')V.$$

Or, les termes constants de e^S, e^T et e^{S+T} sont tous égaux à 1, et il en est de même du terme constant de $e^S e^T$. Ainsi, U et V sont toutes deux solutions de l'équation

$$F' - (S'+T')F = 0, \quad F(0) = 1,$$

et sont donc égales d'après la proposition 3.1. L'application exp est donc bien un morphisme de groupes.

Posons maintenant $V = e^{\ln(1+X)}$. Puisque $\ln(1+X) \in XK[[X]]$, on a par conséquent $V \in 1 + XK[[X]]$, i.e. $V(0) = 1$. De plus, on a

$$(1+X)V' = (1+X)\big((1+X)^{-1}e^{\ln(1+X)}\big) = V.$$

Par ailleurs, le terme constant de $1+X$ est 1, et l'on a aussi

$$(1+X)'(1+X) - (1+X) = 0.$$

Ainsi, V et $1+X$ sont deux séries formelles toutes deux solutions de l'équation différentielle

$$(1+X)F' - F = 0, \quad F(0) = 1.$$

Elles sont donc égales. Autrement dit, on a

$$e^X \circ \ln(1+X) = 1 + X.$$

En composant par $S \in XK[[X]]$, et en utilisant l'associativité de $\circ$, on obtient

$$e^{\ln(1+S)} = 1 + S, \quad \text{pour tout } S \in XK[[X]].$$

Posons enfin $W = \ln(e^X)$. Puisque $e^X \in 1 + XK[[X]]$, on a $W \in XK[[X]]$, soit $W(0) = 0$.

De plus, on a

$$W' = \ln\big(1 + (e^X - 1)\big)' = \big(1 + (e^X - 1)\big)^{-1} e^X = 1.$$

Ainsi, W et X sont toutes deux solutions de l'équation différentielle

$$F' = 1, \quad F(0) = 1.$$

Elles sont donc égales par le lemme 2.16 (1). On a donc

$$\ln(e^X) = X.$$

Par composition avec $S \in XK[[X]]$, on obtient comme précédemment l'égalité

$$\ln(e^S) = S, \quad \text{pour tout } S \in XK[[X]].$$

Ainsi, exp et ln sont des bijections réciproques. Cela achève la démonstration. □

3.3. Remarques

(1) La proposition précédente et les propriétés générales des morphismes de groupes impliquent que l'application

$$\ln : 1 + XK[[X]] \longrightarrow XK[[X]]$$

est un morphisme de groupes. Autrement dit, on a

$$\ln(T_1T_2) = \ln(T_1) + \ln(T_2), \quad \text{pour tous } T_1, T_2 \in 1 + XK[[X]].$$

(2) L'application exp étant un morphisme, on a $e^0 = 1$ ainsi que

$$(e^X)^{-1} = e^{-X}.$$

De plus, pour tout $m \in \mathbb{Z}$, on a $(e^X)^m = e^{mX}$.

Nous allons maintenant généraliser l'exponentiation des séries formelles. Soit $\alpha \in K$. Pour tout $k \geqslant 0$, on pose

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)\times\cdots\times(\alpha-k+1)}{k!},$$

avec la convention $\binom{\alpha}{0} = 1$. On définit alors

$$(1+X)^\alpha = \sum_{n\geqslant 0} \binom{\alpha}{n} X^n.$$

Pour tout $S \in XK[[X]]$, on peut alors définir $(1+S)^\alpha$ par substitution, ou encore T^α pour tout $T \in 1 + XK[[X]]$.

3.4. Remarques

(1) La définition montre que $T^0 = 1$ et $T^1 = T$ pour tout $T \in 1+XK[[X]]$.

(2) Pour tout $T \in 1 + XK[[X]]$, et tout $\alpha \in K$, on a $T^\alpha \in 1 + XK[[X]]$. En effet, si $T = 1 + S$ avec $v_X(S) \geqslant 1$, alors $v_X(\binom{\alpha}{n}S^n) \geqslant n \geqslant 1$ pour tout $n \geqslant 1$, et ainsi $T^\alpha(0) = 1$.

(3) Si $m \geqslant 0$, alors $\binom{m}{k} = 0$ pour tout $k > m$, et coïncide avec le coefficient binomial classique pour tout $k \in [\![0, m]\!]$. On obtient alors

bien le produit $(1+X)^m = (1+X)\cdots(1+X)$ au sens classique, d'où la notation choisie.

On a alors également $T^m = T\cdots T$, pour tout $T \in 1+XK[[X]]$.

On a alors le résultat suivant.

3.5. Proposition. *Pour tous $\alpha, \beta \in K$, et tous $T, T_1, T_2 \in 1+XK[[X]]$, on a*

(1) $(T^\alpha)' = \alpha T' T^{\alpha-1}$;

(2) $T^{\alpha+\beta} = T^\alpha T^\beta$;

(3) $T^{\alpha\beta} = (T^\alpha)^\beta$.

(4) $(T^\alpha)^{-1} = T^{-\alpha} = (T^{-1})^\alpha$;

(5) $(T_1T_2)^\alpha = T_1^\alpha T_2^\alpha$.

Démonstration. Soit $\alpha \in K$. On a alors

$$\big((1+X)^\alpha\big)' = \sum_{n\geqslant 0}(n+1)\binom{\alpha}{n+1}X^n.$$

Mais on a

$$\begin{aligned}(n+1)\binom{\alpha}{n+1} &= (n+1)\frac{\alpha\times(\alpha-1)\times\cdots\times(\alpha-(n+1)+1)}{(n+1)!}\\ &= \alpha\frac{(\alpha-1)\times\cdots\times(\alpha-1-n+1)}{n!}\\ &= \alpha\binom{\alpha-1}{n}.\end{aligned}$$

On en déduit donc

$$\big((1+X)^\alpha\big)' = \alpha\sum_{n\geqslant 0}\binom{\alpha-1}{n}X^n = \alpha(1+X)^{\alpha-1}.$$

Si maintenant $T = 1+S \in 1+XK[[X]]$, on a $S' = T'$, et donc

$$(T^\alpha)' = \big((1+S)^\alpha\big)' = \alpha S'(1+S)^{\alpha-1} = \alpha T' T^{\alpha-1},$$

d'où (1).

Pour démontrer (2) et (3), il suffit de le vérifier lorsque $T = 1+X$, puis d'obtenir le cas général par substitution.

On commence d'abord par un cas particulier, à savoir le cas $\beta = 1$. On doit donc vérifier l'égalité

$$(1+X)^{\alpha+1} = (1+X)^\alpha(1+X), \qquad \text{pour tout } \alpha \in K.$$

En utilisant la définition de $(1+X)^\alpha$ et en regroupant les termes, on voit

que cela revient à démontrer l'égalité

$$\binom{\alpha}{n} + \binom{\alpha}{n-1} = \binom{\alpha+1}{n}, \quad \text{pour tout } n \geqslant 1,$$

ce qui est un exercice facile laissé au lecteur.

Revenons au cas général. Soient α, $\beta \in K$. Posons

$$U = (1+X)^{\alpha+\beta} \quad \text{et} \quad V = (1+X)^{\alpha}(1+X)^{\beta}.$$

On a $U(0) = V(0) = 1$. De plus, on a

$$(1+X)U' - (\alpha+\beta)U = 0$$

d'après le point (1) et le cas précédent. Mais on a aussi

$$V' = \alpha(1+X)^{\alpha-1}(1+X)^{\beta} + \beta(1+X)^{\alpha}(1+X)^{\beta-1},$$

d'où

$$(1+X)V' - (\alpha+\beta)V = 0.$$

On en déduit que $U = V$ en utilisant la proposition 3.1.

Montrons le point (3). On pose cette fois $U = (1+X)^{\alpha\beta}$ et $V = \big((1+X)^{\alpha}\big)^{\beta}$. On a encore $U(0) = V(0) = 1$. On a également $(1+X)U' - \alpha\beta U = 0$. Par ailleurs,

$$V' = \beta(\alpha(1+X)^{\alpha-1})\big((1+X)^{\alpha}\big)^{\beta-1},$$

soit $(1+X)V' = \alpha\beta\big((1+X)^{\alpha}\big)^{\beta}$, en utilisant le point (1). Autrement dit, les séries formelles U et V sont deux solutions de l'équation

$$(1+X)F' - \alpha\beta F = 0, \quad F(0) = 1.$$

On a donc $U = V$, d'où (3). Le point (4) découle alors de (2) et (3).

Il reste à vérifier (5). Soient T_1, $T_2 \in 1 + XK[[X]]$. On pose $U = (T_1T_2)^{\alpha}$ et $V = T_1^{\alpha}T_2^{\alpha}$. On a $U(0) = V(0) = 1$. De plus,

$$U' = \alpha(T_1'T_2 + T_1T_2')(T_1T_2)^{\alpha-1} = \alpha(T_1'T_1^{-1} + T_2'T_2^{-1})(T_1T_2)^{\alpha},$$

en utilisant le point (2), soit

$$U' - \alpha(T_1'T_1^{-1} + T_2'T_2^{-1})U = 0.$$

Cela dit,

$$V' = (\alpha T_1'T_1^{\alpha-1})T_2^{\alpha} + T_1^{\alpha}(\alpha T_2'T_2^{\alpha-1}) = \alpha(T_1'T_1^{-1} + T_2'T_2^{-1})T_1^{\alpha}T_2^{\alpha},$$

soit

$$V' - \alpha(T_1'T_1^{-1} + T_2'T_2^{-1})V = 0.$$

On en déduit alors comme dans les autres cas que $U = V$, et cela achève la démonstration. □

3.6. Remarque. Le point (4) de la proposition précédente montre alors que T^{-1} (au sens de l'exponentiation des séries formelles) est bien l'inverse de T.

3.7. Corollaire. *Pour tout $\alpha \in K^\times$, et tout $T_0 \in 1 + XK[[X]]$, l'équation*

$$T^\alpha = T_0, \quad T \in 1 + XK[[X]]$$

admet une unique solution.

Démonstration. Puisque l'on a $T_0 \in 1 + XK[[X]]$, on peut définir la série formelle $T = T_0^{\frac{1}{\alpha}} \in 1 + XK[[X]]$. D'après la proposition 3.5, on a alors

$$T^\alpha = T_0^1 = T_0.$$

Si maintenant $T_1, T_2 \in 1 + XK[[X]]$ vérifient $T_1^\alpha = T_0 = T_2^\alpha$, on a alors

$$1 = T_1^\alpha (T_2^\alpha)^{-1} = T_1^\alpha (T_2^{-1})^\alpha = (T_1 T_2^{-1})^\alpha,$$

d'après les points (4) et (5) de la proposition 3.5. Posons $U = T_1 T_2^{-1}$. On a donc $U^\alpha = 1$, et par dérivation $\alpha U' U^{\alpha-1} = 0$.

Comme $U^{\alpha-1}$ est inversible (point (4) de la proposition citée plus haut), et que $\alpha \in K^\times$, on obtient $U' = 0$. Comme K est de caractéristique nulle, on en déduit que U est constante, et puisque $U(0) = 1$, on a $U = 1$, soit encore $T_1 = T_2$, d'où le résultat. □

On a alors le théorème suivant.

3.8. Théorème. *Soit $m \geqslant 1$ un entier, soit $F \in K[[X]] \setminus \{0\}$.*

On écrit $F = \sum_{n \geqslant r} a_n X^n$, avec $a_r \in K^\times$. Alors, l'équation $S^m = F$ a au moins une solution si, et seulement, si $m \mid r$ et $a_r \in K^{\times m}$.

Dans ce cas, si S est une solution particulière de l'équation, l'ensemble des solutions de l'équation est

$$\{\lambda S \mid \lambda \in K^\times,\ \lambda^m = 1\}.$$

Démonstration. Notons que toute série formelle $T \in K[[X]]$ non nulle s'écrit

$$T = c_\ell X^\ell T_0,$$

où $\ell = v_X(T)$, c_ℓ est le coefficient de T en X^ℓ et $T_0 \in 1 + XK[[X]]$.

Supposons que $S \in K[[X]]$ vérifie $S^m = F$. Écrivons

$$S = b_k X^k S_0 \quad \text{et} \quad F = a_r X^r F_0,$$

où $S_0, F_0 \in 1 + XK[[X]]$, $k = v_X(S)$, $r = v_X(F)$, b_k est le coefficient de S

en X^k et a_r est le coefficient de F en X^r. On a ainsi

$$b_k^{km} X^{km} S_0^m = a_r X^r F_0.$$

En comparant les ordres, on obtient $mk = r$. En comparant ensuite les coefficients en X^r, on obtient $a_r = (b_k^k)^m \in K^{\times m}$

Inversement, supposons que $m \mid r$ et que $a_r \in K^{\times m}$. Écrivons alors $r = km$, avec $k \geqslant 0$, et $a_r = b^m$, où $b \in K^\times$. On a alors

$$F = b^m X^{km} F_0,$$

avec $F_0 \in 1 + XK[[X]]$. Par la proposition précédente, il existe une série formelle $S_0 \in 1 + XK[[X]]$ tel que $F_0 = S_0^m$. Si l'on pose $S = bX^k S_0$, on a alors $S^m = F$.

Si $S^m = F$ et $\lambda \in K^\times$ vérifie $\lambda^m = 1$, alors

$$(\lambda S)^m = \lambda^m S^m = F.$$

Inversement, soient S, $T \in K[[X]]$ vérifiant $S^m = F = T^m$. Si $S = b_k X^k S_0$ et $T = c_\ell X^\ell T_0$, avec des notations évidentes, une comparaison des ordres montre que $k = \ell$. On a alors $b_k^m = c_k^m$ et $S_0^m = T_0^m$. La proposition précédente montre que $S_0 = T_0$. De plus, si $\lambda = b_k c_k^{-1}$, on a $\lambda^m = 1$, ainsi que $c_k = \lambda b_k$. On a alors

$$T = \lambda b_k X^k S_0 = \lambda S.$$

Cela achève la démonstration. □

On peut définir bien d'autres séries formelles classiques, par analogie avec les développements en série entière des fonctions usuelles, sur un corps K de caractéristique nulle. En voici quelques exemples.

(1) $\cos(X) = \dfrac{e^{iX} + e^{-iX}}{2} = \displaystyle\sum_{n \geqslant 0} \frac{(-1)^n}{(2n)!} X^{2n}$;

(2) $\sin(X) = \dfrac{e^{iX} - e^{-iX}}{2i} = \displaystyle\sum_{n \geqslant 0} \frac{(-1)^n}{(2n+1)!} X^{2n+1}$;

(3) $\mathrm{ch}(X) = \dfrac{e^{X} + e^{-X}}{2} = \displaystyle\sum_{n \geqslant 0} \frac{1}{(2n)!} X^{2n}$;

(4) $\mathrm{sh}(X) = \dfrac{e^{X} - e^{-X}}{2} = \displaystyle\sum_{n \geqslant 0} \frac{1}{(2n+1)!} X^{2n+1}$;

(5) $\tan(X) = \sin(X)\big(\cos(X)\big)^{-1}$, $\mathrm{th}(X) = \mathrm{sh}(X)\big(\mathrm{ch}(X)\big)^{-1}$;

On laisse le soin au lecteur de définir $\mathrm{Arcsin}(X)$, $\mathrm{Argsh}(X)$, $\mathrm{Argth}(X)$, etc.

4. Quelques applications des séries formelles

Commençons par une application simple à la combinatoire.

4.1. Exemple. Pour tout entier $n \geqslant 0$, on note D_n le nombre de permutations $\sigma \in \mathfrak{S}_n$ sans points fixes, avec la convention $D_0 = 1$.

Puisque le support d'une permutation σ est stable par σ et qu'aucun des éléments du support n'est fixe, σ est entièrement déterminée par le choix de son support F, et d'une permutation sans points fixes de F.

Ainsi, pour tout $k \in [\![0, n]\!]$, il y a exactement $\binom{n}{n-k} D_{n-k}$ permutations avec exactement k points fixes. En regroupant les éléments de $\mathfrak{S}_n$ selon leur nombre de points fixes, et en tenant compte de l'égalité $\binom{n}{n-k} = \binom{n}{k}$, on obtient donc

$$\sum_{k=0}^{n} \binom{n}{k} D_{n-k} = n!, \quad \text{pour tout } n \geqslant 0,$$

soit encore

$$\sum_{k=0}^{n} \frac{1}{k!} \times \frac{D_{n-k}}{(n-k)!} = 1, \quad \text{pour tout } n \geqslant 0,$$

Posons $S = \sum_{n \geqslant 0} \frac{D_n}{n!} X^n \in \mathbb{C}[[X]]$. La relation précédente se récrit alors

$$e^X S = \sum_{n \geqslant 0} X^n.$$

En multipliant par e^{-X} les deux membres de l'égalité, on obtient

$$S = e^{-X} \Big(\sum_{n \geqslant 0} X^n \Big) = \sum_{n \geqslant 0} \Big(\sum_{k=0}^{n} \frac{(-1)^k}{k!} \Big) X^n.$$

En identifiant les coefficients, on obtient alors

$$D_n = n! \Big(\sum_{k=0}^{n} \frac{(-1)^k}{k!} \Big) = \sum_{k=0}^{n} (-1)^k \binom{n}{k} (n-k)!, \quad \text{pour tout } n \geqslant 0.$$

Nous continuons par des considérations sur les fractions rationnelles.

Notation. Soit K un corps. On note $K(X)_0$ l'ensemble des fractions rationnelles sans pôle en 0, i.e. les fractions admettant au moins une écriture de la forme $\frac{P}{Q}$, avec $P, Q \in K[X]$, et $Q(0) \neq 0$.

C'est une sous-algèbre de $K(X)$. Remarquons que $Q(0) \in K^\times$, et donc que Q est inversible dans $K[[X]]$.

La démonstration du lemme suivant est laissée au lecteur.

4.2. Lemme. *Soit K un corps, et soit $F = \frac{P}{Q} \in K(X)_0$, avec $Q(0) \neq 0$. Alors, la série formelle PQ^{-1} est indépendante de la représentation de F choisie, et l'application*

$$\begin{aligned} \psi : K(X)_0 &\longrightarrow K[[X]] \\ \frac{P}{Q} &\longmapsto PQ^{-1} \end{aligned}$$

est un morphisme de K-algèbres injectif.

4.3. Définition. Soit $F \in K(X)_0$. L'image de F par le morphisme injectif est appelé le *développement en série formelle (D.S.F.)* de F .

Soit $F \in K(X)_0$ et soit $G \in XK(X)_0$. Écrivons $F = \frac{P}{Q}$, $P, Q \in K[X]$, avec $Q(0) \neq 0$. Considérons le morphisme d'évaluation

$$\begin{aligned} ev_G : K[X] &\longrightarrow K(X) \\ R &\longmapsto R(G). \end{aligned}$$

Si $R = a_d X^d + \cdots + a_1 X + a_0$, on vérifie sans mal que l'on a

$$R(G) = a_d G^d + \cdots + a_1 G + a_0 = a_0 + X\tilde{G},$$

pour un certain $\tilde{G} \in K(X)_0$. En particulier, on a $R(G) \in K(X)_0$. De plus, si $R(0) \neq 0$, alors $R(G)$ est inversible dans $K(X)_0$.

On pose alors

$$F \circ G = \frac{P(G)}{Q(G)} \in K(X)_0.$$

On vérifie aisément que $F \circ G$ ne dépend pas du choix de la représentation de F, et que l'application

$$\begin{aligned} K(X)_0 &\longrightarrow K(X)_0 \\ F &\longmapsto F \circ G \end{aligned}$$

est un morphisme de K-algèbres.

De plus, l'application de dérivation formelle des polynômes s'étend évidemment à $K(X)$ comme suit : si $F = \frac{P}{Q} \in K(X)$, on pose

$$F' = \Big(\frac{P}{Q}\Big)' = \frac{P'Q - PQ'}{Q^2}.$$

On vérifie que cela ne dépend pas de la représentation choisie, et que les propriétés usuelles de dérivation s'étendent à $K(X)$. Remarquons aussi que si $F \in K(X)_0$, il en est de même de F'.

On a alors le résultat suivant.

4.4. Proposition. *Soit $F \in K(X)_0$ et soit $G \in XK(X)_0$. Alors :*

(1) $\psi(F') = \psi(F)'$;

(2) $\psi(F \circ G) = \psi(F) \circ \psi(G)$.

Démonstration. Écrivons $F = \dfrac{P}{Q}$, avec $Q(0) \neq 0$. On a $\psi(F) = PQ^{-1}$. Mais alors,

$$\big(\psi(F)\big)' = P'Q^{-1} + P(-Q'Q^{-2}) = (P'Q - PQ')(Q^2)^{-1}.$$

Or, $F' = \dfrac{P'Q - PQ'}{Q^2}$, et donc

$$\psi(F') = (P'Q - PQ')(Q^2)^{-1} = \big(\psi(F)\big)',$$

d'où (1).

Montrons (2). L'application

$$\begin{aligned} \varphi_G : K(X)_0 &\longrightarrow K[[X]] \\ F &\longmapsto \psi(F \circ G) \end{aligned}$$

est un morphisme de K-algèbres, comme application composée de deux morphismes de K-algèbres.

Soit $R = \sum_{n \geqslant 0} a_n X^n \in K[X]$ un polynôme. On a alors

$$\psi(R \circ G) = \varphi_G(R) = \sum_{n \geqslant 0} a_n \varphi_G(X)^n = \sum_{n \geqslant 0} a_n \psi(G)^n = R \circ \psi(G) \in K[[X]].$$

Comme $P = QF$, on a $\varphi_G(P) = \varphi_G(Q)\varphi_G(F)$.

En tenant compte de ce qui précède, on obtient

$$P \circ \psi(G) = \big(Q \circ \psi(G)\big)\psi(F \circ G).$$

Puisque $G \in XK(X)_0$, et puisque ψ est un morphisme de K-algèbres, on a alors $\psi(G) \in XK[[X]]$, soit encore $v_X\big(\psi(G)\big) \geqslant 1$. Or, on a $Q(0) \neq 0$ et donc $v_X(Q) = 0$. Par conséquent,

$$v_X\big(Q \circ \psi(G)\big) = v_X(Q) v_X\big(\psi(G)\big) = 0,$$

et $Q \circ \psi(G) \in K[[X]]^\times$. On a alors

$$\psi(F \circ G) = \big(P \circ \psi(G)\big)\big(Q \circ \psi(G)\big)^{-1} = (PQ^{-1}) \circ \psi(G) = \psi(F) \circ \psi(G).$$

Cela achève la démonstration. □

Les résultats précédents montrent que l'on peut identifier une fraction rationnelle sans pôle en 0 à son D.S.F., et que cette identification est com-

patible avec la somme, le produit, la dérivation et la substitution. En particulier, si $Q \in K[X]$ est telle que $Q(0) \neq 0$, cette identification permet de confondre $Q^{-1} \in K[[X]]$ avec $\dfrac{1}{Q}$ (ou plutôt son D.S.F.), ce qui permet par exemple d'écrire sans remords

$$\frac{1}{1-X} = \sum_{n \geqslant 0} X^n.$$

4.5. Exemple. Le D.S.F. de $\dfrac{1}{1-X}$ est bien entendu $\sum\limits_{n \geqslant 0} X^n$. En dérivant d fois successivement, on obtient

$$\frac{d!}{(1-X)^{d+1}} = \sum_{n \geqslant d} n \times (n-1) \times \cdots \times (n-d+1) X^{n-d}.$$

La théorie des séries formelles permet de retrouver l'expression des polynômes de Tchebychev de seconde espèce.

4.6. Exemple. Soit $\theta \in \mathbb{R} \setminus \pi\mathbb{Z}$ et soit $F = \dfrac{1}{X^2 - 2\cos(\theta)X + 1}$.
On a

$$F = \frac{1}{1 - \left(2\cos(\theta)X - X^2\right)} = \sum_{n \geqslant 0} \left(2\cos(\theta)X - X^2\right)^n,$$

soit encore

$$F = \sum_{n \geqslant 0} \left(\sum_{k=0}^{n} (-1)^k \binom{n}{k} 2^{n-k} \cos(\theta)^{n-k} X^{n+k} \right).$$

On a donc $F = \sum\limits_{m \geqslant 0} u_m(\theta) X^m$, avec

$$u_m(\theta) = \sum_{n+k=m, k \in [\![0,n]\!]} (-1)^k \binom{n}{k} 2^{n-k} \cos(\theta)^{n-k},$$

ou encore

$$u_m(\theta) = \sum_{0 \leqslant k \leqslant \frac{m}{2}} (-1)^k \binom{m-k}{k} 2^{m-2k} \cos(\theta)^{m-2k}.$$

Par ailleurs, on a aussi

$$F = \frac{1}{(1 - Xe^{-i\theta})(1 - Xe^{i\theta})} = \frac{i}{2\sin(\theta)} \left(\frac{e^{-i\theta}}{1 - Xe^{-i\theta}} - \frac{e^{i\theta}}{1 - Xe^{i\theta}} \right).$$

Des substitutions appropriées montrent alors que

$$F = \frac{i}{2\sin(\theta)} \sum_{m \geqslant 0} (e^{-i(m+1)\theta} - e^{i(m+1)\theta}) X^m = \sum_{m \geqslant 0} \frac{\sin(m+1)\theta}{\sin(\theta)} X^m,$$

d'où l'identité bien connue

$$\frac{\sin(m+1)\theta}{\sin(\theta)} = \sum_{0\leqslant k\leqslant \frac{m}{2}} (-1)^k \binom{m-k}{k} 2^{m-2k} \cos(\theta)^{m-2k},$$

pour tout $m \geqslant 0$.

Continuons maintenant par une autre application combinatoire.

4.7. Exemple. Soient $a_1, \ldots, a_p$ des entiers $\geqslant 1$ deux à deux distincts, et considérons

$$F = \frac{1}{\prod_{k=1}^{p} (1 - X^{a_k})}.$$

On a alors

$$F = \prod_{k=1}^{p} \Big(\sum_{n\geqslant 0} X^{na_k} \Big) = \sum_{n\geqslant 0} u_n X^n,$$

où $u_n = \big|\{(k_1, \ldots, k_p) \in \mathbb{N}^p \mid k_1 a_1 + \cdots + k_p a_p = n\}\big|$.

En décomposant F en éléments simples sur $\mathbb{C}$ par les méthodes usuelles, et en utilisant l'égalité

$$\frac{d!}{(1-zX)^{d+1}} = \sum_{n\geqslant d} n\times(n-1)\times\cdots\times(n-d+1) z^{n-d} X^{n-d},$$

pour tout $d \geqslant 0$, on trouve une expression explicite de u_n. À titre d'exercice, le lecteur pourra traiter le cas $p = 2$, $a_1 = 2$, $a_2 = 3$ (cf. exercice 8).

Donnons maintenant deux applications de la formule de réversion de Lagrange.

4.8. Exemple. La série formelle $S = Xe^{-X} \in \mathbb{C}[[X]]$ étant d'ordre 1, elle admet une série réverse. Comme $(e^{-X})^{-k} = e^{kX}$ pour tout $k \geqslant 0$, la formule de réversion de Lagrange donne

$$S^{\langle -1\rangle} = \sum_{n\geqslant 1} \frac{n^{n-1}}{n!} X^n.$$

En évaluant l'identité $Se^X = X$ en $S^{\langle -1\rangle}$ et en utilisant le fait que l'évaluation est un morphisme d'anneaux, on obtient

$$(S \circ S^{\langle -1\rangle}) e^{S^{\langle -1\rangle}} = S^{\langle -1\rangle}.$$

L'égalité $S \circ S^{\langle -1 \rangle} = X$ implique alors l'identité

$$e^{\sum_{n \geqslant 1} \frac{n^{n-1}}{n!} X^n} = \sum_{n \geqslant 1} \frac{n^{n-1}}{n!} X^{n-1}.$$

4.9. Exemple. Pour tout $n \geqslant 0$, considérons le nombre C_n de façons de parenthéser un produit de $n+1$ facteurs (pour une loi interne non associative, par exemple). Il s'agit donc de placer des parenthèses autour des termes de sorte que l'on n'ait jamais plus de deux termes non parenthésés, et tel que l'on ne mette jamais des parenthèses autour d'un seul terme. Bien sûr, $C_0 = 1$.

De plus, en considérant les parenthésages contenant le dernier terme, on voit que se donner un parenthésage de $n+2$ termes revient à se donner un parenthésage de $k+1$ termes et un parenthésage de $(n+2)-(k+1)$ termes, où $k \in [\![0, n]\!]$. On a donc

$$C_{n+1} = \sum_{k=0}^{n} C_k C_{n-k}.$$

Posons $F = \sum_{n \geqslant 0} C_n X^n$. D'après la relation de récurrence précédente, on a alors $F^2 = \sum_{n \geqslant 0} C_{n+1} X^n$, d'où

$$XF^2 = \sum_{n \geqslant 0} C_{n+1} X^{n+1} = \sum_{n \geqslant 1} C_n X^n = F - 1.$$

Posons $W = F - 1$. On a alors $X = W(1+W)^{-2}$. On a donc $W = S^{\langle -1 \rangle}$, avec $S = X(1+X)^{-2}$. On a

$$\left((1+X)^{-2}\right)^{-k} = (1+X)^{2k}, \quad \text{pour tout } k \geqslant 0.$$

La formule de réversion de Lagrange donne alors

$$\sum_{n \geqslant 1} C_n X^n = S^{\langle -1 \rangle} = \sum_{n \geqslant 1} \frac{1}{n} \binom{2n}{n-1} X^n.$$

Pour tout $n \geqslant 1$, on a donc

$$C_n = \frac{1}{n} \binom{2n}{n-1} = \frac{1}{n} \times \frac{(2n)!}{(n+1)! \times (n-1)!},$$

soit encore

$$C_n = \frac{1}{n+1} \binom{2n}{n} \quad \text{pour tout } n \geqslant 1,$$

égalité encore valable pour $n \geqslant 0$.

5. Exercices

E1. Soit A un anneau commutatif.

a. Montrer que A est noethérien si, et seulement si, $A[[X]]$ est noethérien.

Indication. Imiter la démonstration du théorème XXII-2.2, en considérant v_X à la place du degré.

b. Montrer l'équivalence des propriétés suivantes :

(i) $A[[X]]$ est euclidien ;

(ii) $A[[X]]$ est principal ;

(iii) $A[[X]]$ est un corps.

E2.

a. Soit A un anneau principal, et soit $S \in A[[X]]$ une série formelle irréductible telle que $S(0) \neq 0$. Montrer par l'absurde que $S(0)$ est associé à une puissance non triviale d'un élément premier de A.

Indication. Montrer que $S(0)$ est non inversible. Si $S(0)$ n'est pas associé à une puissance non triviale d'un élément premier, alors $S(0) = ab$, avec a, $b \in A$ non nuls, non inversibles, et premiers entre eux.

Soient alors α, $\beta \in A$ tels que $\alpha a + \beta b = 1$. Montrer alors que l'on peut écrire $S = S_1 S_2$, avec $S_1(0) = a$ et $S_2(0) = b$.

b. Le polynôme $3X^2 + 7X + 2$ est-il irréductible dans $\mathbb{Z}[X]$? dans $\mathbb{Z}[[X]]$?

c. Soit $P = X^2 + X + 6 \in \mathbb{Z}[X]$.

(i) Le polynôme P est-il irréductible dans $\mathbb{Z}[X]$? dans $\mathbb{Z}[[X]]$?

(ii) Le polynôme P a-t-il une racine dans $\mathbb{Z}[[X]]$?

(iii) Justifier que la série formelle $S = \dfrac{1 - \sqrt{1 - 4(X^2 + X)}}{2}$ est bien définie et appartient à $\mathbb{Z}[[X]]$, que $2 + S$ et $3 - S$ ne sont pas inversibles, et que l'on a l'égalité

$$X^2 + X + 6 = (2 + S)(3 - S).$$

E3. Soit A un anneau principal.

a. Soit $\mathfrak{a}$ un idéal de $A[[X]]$ contenant X. Montrer qu'il existe $d \in A$ tel que $\mathfrak{a} = (d, X)$.

Indication. Montrer que si $S \in \mathfrak{a}$, on a $S(0) \in \mathfrak{a}$.

b. En déduire que les idéaux premiers de $A[[X]]$ contenant X sont (X) et les idéaux (π, X), où π est un élément irréductible de A.

E4. Soit K un corps.

Montrer que les automorphismes de K-algèbres de $K[[X]]$ sont exactement les morphismes de la forme

$$\begin{aligned} K[[X]] &\longrightarrow K[[X]] \\ F &\longmapsto F \circ S, \end{aligned}$$

avec $v_X(S) = 1$.

E5. Soit K un corps, et soient $A, B \in K[[X]]$.

a. On suppose que $\mathrm{car}(K) \neq 2$. Donner une condition nécessaire et suffisante sur A et B pour que l'équation

$$S^2 + AS + B = 0, \quad S \in K[[X]]$$

ait au moins une solution.

b. On suppose que $\mathrm{car}(K) = 2$. Montrer que l'équation $S^2 + S + B = 0$ a une solution si, et seulement si, il existe $a_0 \in K$ tel que $B(0) = a_0^2 + a_0$.

c. Résoudre l'équation $S^2 + S + X = 0$ dans $K[[X]]$, selon que K est de caractéristique différente de 2 ou non.

E6. Soit K un corps.

a. Soit $d \geqslant 0$ un entier. Vérifier que la famille

$$1,\ X,\ X(X-1),\ \ldots,\ X(X-1)\cdots(X-d+1)$$

est une K-base de l'espace vectoriel des polynômes de degré $\leqslant d$.

Soit $P \in K[X]$ de degré $d \geqslant 0$, que l'on écrit

$$P = a_d X(X-1)\cdots(X-d+1) + \cdots + a_1 X + a_0, \quad a_0, \ldots, a_d \in K[X].$$

b. Calculer $\sum_{n \geqslant 0} P(n) X^n$ en termes de séries formelles usuelles.

c. Calculer $\sum_{n \geqslant 0} \frac{P(n)}{n!} X^n$ en termes de séries formelles usuelles, lorsque le corps K est de caractéristique nulle.

E7. Démontrer le lemme 4.2.

E8.

a. Décomposer la fraction rationnelle $F = \dfrac{1}{(1-X^2)(1-X^3)}$ en éléments simples dans $\mathbb{R}[X]$.

b. En déduire le D.S.F. de F dans $\mathbb{R}[[X]]$. Pour faciliter les calculs, on pourra remarquer que $\dfrac{1}{X^2+X+1} = \dfrac{1-X}{1-X^3}$.

c. Pour tout $n \geqslant 0$, en déduire $u_n = \big|\{(a,b) \in \mathbb{N}^2 \mid 2a+3b=n\}\big|$. On distinguera les cas selon la classe de congruence de n modulo 6.

E9. Soit K un corps. On rappelle que $K(X)_0$ désigne la K-algèbre des fractions rationnelles sans pôle en 0.

a. Soit $S = \sum_{n\geqslant 0} a_n X^n \in K[[X]]$. Montrer que S est le D.S.F. d'une fraction rationnelle $F \in K(X)_0$ si, et seulement si, il existe $\ell \geqslant 1$, $m \geqslant 0$ et $\lambda_1, \ldots, \lambda_\ell \in K$, avec $\lambda_\ell \in K^\times$, vérifiant

$$a_{n+\ell} = -\lambda_1 a_{n+\ell-1} - \cdots - \lambda_\ell a_n, \quad \text{pour tout } n \geqslant m.$$

Montrer que dans ce cas, si ℓ est minimal, alors le polynôme

$$\lambda_\ell X^\ell + \cdots + \lambda_1 X + 1$$

est le dénominateur d'une représentation irréductible de F.

Dans la suite, on suppose que $S = \sum_{n\geqslant 0} a_n X^n \in K[[X]] \setminus K[X]$. Pour tout $n \geqslant 0$, on définit la matrice

$$\mathscr{H}_n = (a_{i+j})_{i,j\in[\![0,n]\!]} \in \mathrm{M}_{n+1}(K).$$

On veut montrer que S est le D.S.F. d'un élément de $K(X)_0$ si, et seulement si, il existe $p \geqslant 1$ tel que $\det(\mathscr{H}_n) = 0$ pour tout $n \geqslant p$.

b. Montrer la nécessité de la condition.

c. On suppose il existe $p \geqslant 1$ tel que $\det(\mathscr{H}_n) = 0$ pour tout $n \geqslant p$. Pour tout $n \geqslant 0$, notons $\Delta_n = \det(\mathscr{H}_n)$.

(i) Soit $r = v_X(S)$. Montrer que $\Delta_r \neq 0$. En déduire qu'il existe $s \geqslant 0$ tel que $\Delta_s \neq 0$ et $\Delta_n = 0$ pour tout $n > s$.

(ii) Montrer que la dernière colonne de $\mathscr{H}_{s+1}$ est combinaison linéaire des autres colonnes. En déduire l'existence de $\mu_0, \ldots, \mu_s \in K$ tels que

$$a_n = \mu_0 a_{n-s-1} + \cdots + \mu_s a_{n-1}, \quad \text{pour tout } n \in [\![s+1, 2s+2]\!].$$

(iii) Soit $k \geqslant 0$. Par des opérations élémentaires sur les colonnes de la matrice $\mathscr{H}_{s+k+2}$, montrer qu'il existe $\alpha \in K$ tel que

$$\Delta_{s+k+2} = \Delta_s (-1)^{k+1} \alpha^{k+2}.$$

(iv) Justifier que $\alpha = 0$. En déduire par récurrence que l'on a

$$a_n = \mu_0 a_{n-s-1} + \cdots + \mu_s a_{n-1}, \quad \text{pour tout } n \geqslant s+1,$$

et conclure [2].

2. Cet exercice est tiré de la démonstration du théorème V.17 de [1].

Les quatre exercices suivants sont tirés de [3].

E10. On considère la suite de Fibonacci $(F_n)_{n\geqslant 0}$ définie par

$$F_{n+2} = F_{n+1} + F_n, \quad \text{pour tout } n \geqslant 0, \ F_0 = F_1 = 1.$$

On pose $S = \sum_{n\geqslant 0} F_n X^n \in \mathbb{C}[[X]]$.

a. Montrer que $S = \dfrac{1}{1 - X - X^2}$, puis en décomposant S en éléments simples, en déduire que

$$F_n = \frac{1}{\sqrt{5}}\left(\omega^{n+1} - (1-\omega)^{n+1}\right), \quad \text{pour tout } n \geqslant 0,$$

où $\omega = \dfrac{1+\sqrt{5}}{2}$.

b. Soit $T = \sum_{n\geqslant 0} F_n^2 X^n$. Montrer que $T = \dfrac{1 - X}{1 - 2X - 2X^2 + X^3}$.

c. Montrer que $F_n^2 + F_{n+1}^2 = F_{2n+2}$ pour tout $n \geqslant 0$.

E11. On rappelle qu'une partition d'un ensemble non vide E est un ensemble de parties non vides de E, deux à deux disjointes, et de réunion égale à E.

Pour tout $n \geqslant 0$, on note π_n le nombre de partitions de $E_n = [\![1, n]\!]$. Par convention, on convient que $\pi_0 = 1$.

a. Montrer que $\pi_{n+1} = \sum_{k=0}^{n} \binom{n}{k} \pi_{n-k}$, pour tout $n \geqslant 1$.

Indication. Classer les partitions de E_{n+1} selon le cardinal de l'unique élément de la partition contenant $n+1$.

On pose $S = \sum_{n\geqslant 0} \dfrac{\pi_n}{n!} X^n \in \mathbb{Q}[[X]]$.

b. Montrer que $S' = e^X S$ et $S(0) = 1$. En considérant $T = e^{-X} S$, en déduire que $S = e^{e^X - 1}$.

La série formelle étant d'ordre 1, $(e^X - 1)^k$ est d'ordre k. Pour tout entier $k \geqslant 0$, on pose donc

$$\frac{(e^X - 1)^k}{k!} = \sum_{n\geqslant k} \frac{S_{n,k}}{n!} X^n \in \mathbb{Q}[[X]].$$

c. Vérifier que $S_{0,0} = 1$, $S_{n,0} = 0$ et $\pi_n = \sum_{k=0}^{n} S_{n,k}$, pour tout $n \geqslant 0$.

d. Montrer que l'on a $S_{n,k} = \dfrac{1}{k!}\displaystyle\sum_{j=0}^{k}(-1)^{k-j}\binom{k}{j}j^n$, pour tous $n \geqslant k \geqslant 0$.

e. En remarquant que $\Big(\dfrac{(e^X-1)^k}{k!}\Big)' = e^X\,\dfrac{(e^X-1)^{k-1}}{(k-1)!}$, pour tout $k \geqslant 1$, montrer que

$$S_{n+1,k} = \sum_{j=k-1}^{n}\binom{n}{j}S_{j,k-1}, \qquad \text{pour tous } n \geqslant k-1 \geqslant 0.$$

f. On définit une suite double d'entiers $(S^*_{n,k})_{n\geqslant k\geqslant 0}$ comme suit.
On pose $S^*_{0,0} = 1$, $S^*_{n,0} = 0$ pour tout $n \geqslant 1$, et pour tout $n \geqslant k \geqslant 1$, on définit $S^*_{n,k}$ comme étant le nombre de partitions de E_n en k sous-ensembles.

En raisonnant comme pour les π_n, montrer que l'on a

$$S^*_{n+1,k} = \sum_{j=k-1}^{n}\binom{n}{j}S^*_{j,k-1}, \qquad \text{pour tous } n \geqslant k-1 \geqslant 0.$$

g. En déduire que pour tous $n \geqslant k \geqslant 1$, on a $S_{n,k} = S^*_{n,k}$.

E12. Par abus de notation, on note $\dfrac{e^X-1}{X} \in \mathbb{C}[[X]]$ l'unique série formelle S telle que $e^X - 1 = XS \in \mathbb{C}[[X]]$. On définit de manière similaire les séries formelles $\dfrac{\tan(X)}{X}$ et $\dfrac{\mathrm{th}(X)}{X}$.

a. Justifier que les séries formelles $\dfrac{e^X-1}{X}$, $\dfrac{\tan(X)}{X}$ et $\dfrac{\mathrm{th}(X)}{X}$ sont inversibles.
On note alors $(b_n)_{n\geqslant 0}$ l'unique suite de nombres complexes telle que

$$\frac{X}{e^X-1} = \sum_{n\geqslant 0}\frac{b_n}{n!}X^n.$$

Le complexe b_n est appelé le n-ième nombre de Bernoulli.

b. Justifier que $b_n \in \mathbb{Q}$, pour tout $n \geqslant 0$.

c. Écrire la relation de récurrence permettant de calculer b_n en fonction des termes précédents, et en déduire l'inégalité

$$|b_n| \leqslant n!, \qquad \text{pour tout } n \geqslant 0.$$

d. Montrer que $\dfrac{X}{e^X-1} + \dfrac{X}{2} = \dfrac{X}{2\mathrm{th}\big(\frac{X}{2}\big)}$, et en déduire que $b_{2n+1} = 0$, pour tout $n \geqslant 1$.

e. En déduire que $\dfrac{X}{\tan(X)} = 1 + \sum_{n\geqslant 1}(-1)^n b_{2n}\dfrac{2^{2n}}{(2n)!}X^{2n}$.

f. Vérifier l'égalité $X\tan(X) = \dfrac{X}{\tan(X)} - \dfrac{2X}{\tan(2X)}$, et en déduire que

$$\tan(X) = \sum_{n\geqslant 1}(-1)^{n+1}\frac{2^{2n}}{(2n)!}(2^{2n}-1)b_{2n}X^{2n-1}.$$

E13. Soit K un corps de caractéristique nulle. Pour tous A, $B \in K[[X]]$ et tout $s_0 \in K$, montrer qu'il existe une unique série formelle $S \in K[[X]]$ telle que

$$S' - AS = B \quad \text{et} \quad S(0) = s_0,$$

et écrire explicitement cette solution en termes de séries formelles usuelles.

Pour toute série formelle $F \in K[[X]]$, on pourra noter $I(F) \in K[[X]]$ l'unique série formelle $G \in XK[[X]]$ vérifiant $G' = F$.

E14. Soit K un corps de caractéristique nulle, et soit $\alpha \in K$.

a. Pour tout $T \in 1 + XK[[X]]$, montrer que $T^\alpha = e^{\alpha\ln(T)}$.

b. Pour tout $S \in XK[[X]]$, montrer que $(e^S)^\alpha = e^{\alpha S}$.

c. Redémontrer les points (2) et (4) de la proposition 3.5 à partir de la relation

$$e^{\alpha ln(T)} = T^\alpha.$$

Chapitre XXIV

Complétion I-adique

Le but de ce chapitre est de réinterpréter la construction et les propriétés de l'anneau des séries formelles de manière plus conceptuelle, et de les généraliser. On suppose le lecteur familier avec les notions de base de topologie générale, en particulier avec les propriétés des espaces métriques. Nous renvoyons le lecteur aux paragraphes 10-15 de la partie « Vocabulaire mathématique » de [20] pour les définitions et résultats de topologie utilisés dans ce chapitre.

Tous les anneaux sont supposés ici commutatifs.

1. Complété I-adique

Nous allons commencer par introduire la notion de complété I-adique par rapport à un idéal. Avant de nous lancer dans une définition précise, commençons par quelques remarques informelles sur les entiers et les polynômes.

Si $m \geqslant 0$ est un entier, il est uniquement déterminé par la connaissance de ses réductions modulo les puissances de 10. En effet, pour tout $k \geqslant 0$, l'unique représentant de m modulo 10^{k+1} compris entre 0 et $10^{k+1}-1$ donnera les k derniers chiffres de m. Remarquons par ailleurs que si l'écriture décimale de m comporte au plus $n+1$ chiffres (i.e. si $m \in [\![0, 10^{n+1}-1]\!]$), la suite des représentants des réductions successives de m modulo les puissances de 10 stationne à partir du rang $n+1$.

Le même phénomène peut être constaté pour les polynômes. En effet, un polynôme $P \in A[X]$ est uniquement déterminé par la connaissance de ses réductions modulo les puissances de X.

De plus, si l'on représente la classe de P modulo X^{k+1} par un polynôme de degré $\leqslant k$, alors la suite des représentants successifs stationne à partir du rang $d = \deg(P)$.

Les choses se passent un peu différemment en ce qui concerne les séries formelles. Une série formelle $S \in A[[X]]$ est encore déterminée de manière unique par ses réductions modulo les puissances de X. En revanche, la suite des représentants successifs n'a absolument aucune raison de stationner en général.

Remarquons que dans les trois cas ci-dessus, les diverses réductions des objets sont compatibles entre elles. Par exemple, si $P_k \in A[X]$ est un représentant de la classe de S modulo X^{k+1}, alors pour tout entier $n \in [\![0, k]\!]$, on a $P_k - P_n \in X^{n+1}A[X]$, ce que l'on notera

$$P_k \equiv P_n \ \ [X^{n+1}].$$

De manière un peu vague, on peut donc voir une série formelle comme une famille d'éléments $(x_k)_{k\geqslant 0}$ avec $x_k \in A[X]/(X^{k+1})$, où les x_k vérifient certaines relations de compatibilité. La classe x_k n'est d'ailleurs rien d'autre que la réduction modulo X^{k+1} de la série formelle initiale.

Plus rigoureusement, on a le résultat suivant.

1.1. Lemme. *Soit A un anneau commutatif, et soit $A[X]_X$ le sous-anneau de $\prod_{n\geqslant 0} A[X]/(X^{n+1})$ défini par*

$$A[X]_X = \{([P_n]_n)_{n\geqslant 0} \mid P_n \equiv P_m \ [X^{m+1}], \quad \textit{pour tous } n \geqslant m \geqslant 0\},$$

où $[P]_n$ désigne la classe de $P \in A[X]$ modulo X^{n+1}.

Alors, l'application

$$\begin{aligned} \varphi : \ A[[X]] &\longrightarrow A[X]_X \\ \sum_{n\geqslant 0} a_n X^n &\longmapsto \Big(\Big[\sum_{k=0}^{n} a_k X^k\Big]_n\Big)_{n\geqslant 0} \end{aligned}$$

est un isomorphisme d'anneaux.

Démonstration. On vérifie aisément que l'application φ prend bien ses valeurs dans $A[X]_X$, et que c'est un morphisme d'anneaux. Montrons que φ est bijective. Soit $\sum_{n\geqslant 0} a_n X^n$ une série formelle telle que

$$\Big[\sum_{k=0}^{n} a_k X^k\Big]_n = [0]_n \in A[X]/(X^{n+1}), \quad \text{pour tout } n \geqslant 0.$$

Comme le polynôme $\sum_{k=0}^{n} a_k X^k$ est de degré au plus n, et qu'un élément non nul de (X^{n+1}) est de degré [1] au moins $n+1$, on en déduit que

$$\sum_{k=0}^{n} a_k X^k = 0, \quad \text{pour tout } n \geqslant 0,$$

d'où $a_n = 0$ pour tout $n \geqslant 0$. Ainsi, φ est injective.

Montrons la surjectivité de φ. Soit $([P_n]_n)_{n\geqslant 0} \in A[X]_X$. On peut toujours supposer que $\deg(P_n) \leqslant n$ pour tout $n \geqslant 0$. Par définition de $A[X]_X$, on a en particulier

$$P_{n+1} \equiv P_n \ [X^{n+1}], \quad \text{pour tout } n \geqslant 0.$$

Écrivons $P_{n+1} = a_{n+1}X^{n+1} + Q_n$, avec $\deg(Q_n) \leqslant n$. On a alors

$$Q_n \equiv P_n \ [X^{n+1}],$$

et comme P_n et Q_n sont de degré au plus n, on en déduit que $Q_n = P_n$. Autrement dit, pour tout $n \geqslant 0$, il existe $a_{n+1} \in A$ tel que

$$P_{n+1} = a_{n+1}X^{n+1} + P_n.$$

En posant $P_0 = a_0$, on obtient alors finalement l'égalité

$$([P_n]_n)_{n\geqslant 0} = \Big(\Big[\sum_{k=0}^{n} a_k X^k\Big]_n\Big)_{n\geqslant 0} = \varphi\Big(\sum_{n\geqslant 0} a_n X^n\Big),$$

d'où la surjectivité. Cela achève la démonstration. □

L'anneau $A[X]_X$ s'appelle le complété de l'anneau $A[X]$ par rapport à l'idéal (X). Le lemme précédent montre alors que l'on obtient l'anneau des séries formelles $A[[X]]$. Nous allons maintenant généraliser ces considérations, afin de l'appliquer à d'autres anneaux pour lesquels un anneau analogue n'a pas été construit, comme l'anneau des entiers par exemple.

Dans tout ce qui suit, A est un anneau commutatif, et I est un idéal.

Notation. Pour tout $n \geqslant 1$, et pour tout $a \in A$, on note $[a]_n$ la classe de a modulo I^{n+1}. Pour tous $n \geqslant m \geqslant 0$, la projection canonique $A \longrightarrow A/I^{m+1}$ induit par passage au quotient un morphisme d'anneaux

$$\begin{aligned} \pi_{n,m} : A/I^{n+1} &\longrightarrow A/I^{m+1} \\ [a]_n &\longmapsto [a]_m. \end{aligned}$$

1. En effet, la définition du degré d'un polynôme montre immédiatement que l'on a l'égalité $\deg(X^{n+1}P) = (n+1) + \deg(P)$, pour tout $P \in A[X]$, même si A n'est pas intègre.

1.2. Définition. Une *suite cohérente* est un élément $(x_n)_{n\geqslant 0} \in \prod_{n\geqslant 0} A/I^{n+1}$ tel que

$$\pi_{n,m}(x_n) = x_m, \quad \text{pour tous } n \geqslant m \geqslant 0.$$

Remarquons que, pour tous $n \geqslant m \geqslant 0$, on a

$$\pi_{n,m} = \pi_{m+1,m} \circ \pi_{m+2,m+1} \circ \cdots \circ \pi_{n,n-1}.$$

On en déduit facilement que $(x_n)_{n\geqslant 0}$ est cohérente si, et seulement si,

$$\pi_{n+1}(x_{n+1}) = x_n, \quad \text{pour tout } n \geqslant 0.$$

L'ensemble A_I des suites cohérentes est un sous-anneau de $\prod_{n\geqslant 0} A/I^{n+1}$, appelé *complété I-adique de A*.

On a donc

$$\begin{aligned} A_I &= \{(x_n)_{n\geqslant 0} \in \prod_{n\geqslant 0} A/I^{n+1} \mid \pi_{n,m}(x_n) = x_m, \text{ pour tous } n \geqslant m \geqslant 0\} \\ &= \{(x_n)_{n\geqslant 0} \in \prod_{n\geqslant 0} A/I^{n+1} \mid \pi_{n+1,n}(x_{n+1}) = x_n, \text{ pour tout } n \geqslant 0\} \\ &= \{([a_n]_n)_{n\geqslant 0} \in \prod_{n\geqslant 0} A/I^{n+1} \mid a_n - a_m \in I^{m+1}, \text{ pour tous } n \geqslant m \geqslant 0\} \\ &= \{([a_n]_n)_{n\geqslant 0} \in \prod_{n\geqslant 0} A/I^{n+1} \mid a_{n+1} - a_n \in I^{n+1}, \text{ pour tout } n \geqslant 0\}. \end{aligned}$$

1.3. Remarque. Soit $(x_n)_{n\geqslant 0} \in A_I$. S'il existe $n \geqslant 0$ tel que $x_n = [0]_n$, alors pour tout $m \in [\![0, n]\!]$, on a

$$x_m = \pi_{n,m}(x_n) = \pi_{n,m}([0]_n) = [0]_m.$$

1.4. Exemples

(1) Pour tout $a \in A$, la suite $([a]_n)_{n\geqslant 0}$ est clairement cohérente.

(2) Pour toute suite $(u_n)_{n\geqslant 0}$ d'éléments de A, et tout $\pi \in I$, la suite

$$\left(\left[\sum_{k=0}^{n} u_k \pi^k\right]_n\right)_{n\geqslant 0}$$

est cohérente.

1.5. Remarque. L'application

$$\begin{aligned} \iota_I : A &\longrightarrow A_I \\ a &\longmapsto ([a]_n)_{n\geqslant 0} \end{aligned}$$

est un morphisme d'anneaux, de noyau $\bigcap_{n\geqslant 0} I^{n+1}$.

Notons que si $I = (0)$, ce morphisme est clairement bijectif. Ainsi, on a un isomorphisme d'anneaux $A_{(0)} \simeq A$.

Notons que pour tout $n \geqslant 0$, on a un morphisme de projection

$$\begin{aligned} \pi_{I,n} : \quad A_I \quad &\longrightarrow A/I^{n+1} \\ x = (x_n)_{n\geqslant 0} &\longmapsto \quad x_n, \end{aligned}$$

compatible avec les diverses projections $\pi_{n,m} : A/I^{n+1} \longrightarrow A/I^{m+1}$. Plus précisément, pour tous $n \geqslant m \geqslant 0$, on a

$$\pi_{n,m} \circ \pi_{I,n} = \pi_{I,m}.$$

Cela provient des définitions des divers morphismes et de la définition d'une suite cohérente.

L'élément $x_n = \pi_{I,n}(x)$ peut donc s'interpréter comme « une approximation de x modulo I^{n+1} », et un élément $x \in A_I$ est donc défini à partir d'approximations successives de plus en plus fines modulo les puissances de l'idéal I.

1.6. Définition. On dit que $x = (x_n)_{n\geqslant 0}$, $y = (y_n)_{n\geqslant 0} \in A_I$ sont *congrus* modulo I^{n+1} si $\pi_{I,n}(x) = \pi_{I,n}(y)$, i.e. si $x_n = y_n \in I^{n+1}$. On le note

$$x \equiv y \mod I^{n+1}.$$

Lorsque I est principal, engendré par un élément $\pi \in I$, on a $I^{n+1} = (\pi^{n+1})$, on le note également

$$x \equiv y \ [\pi^{n+1}].$$

Remarquons aussi que si $x \equiv y \mod I^{n+1}$, on a aussi $x \equiv y \mod I^{m+1}$ pour tout $m \in [\![0, n]\!]$.

Enfin, on dit que $x \in A_I$ est *inversible* modulo I^{n+1} si $\pi_{I,n}(x) = x_n$ est inversible dans l'anneau A/I^{n+1}.

1.7. Remarque. Pour tout $n \geqslant 0$, la projection $\pi_{I,n} : A_I \longrightarrow A/I^{n+1}$ est surjective. En effet, si $[a]_n \in A/I^{n+1}$, on a $\pi_{I,n}(\iota_I(a)) = [a]_n$.

On a déjà vu que l'anneau des séries formelles pouvait s'identifier au complété (X)-adique de $A[X]$. Un autre exemple classique est celui que l'on obtient en complétant $\mathbb{Z}$ par rapport à un idéal non trivial. Cet exemple est suffisamment important en mathématiques pour qu'on lui donne un nom particulier.

1.8. Définition. Soit $b \geqslant 2$ un entier. Le complété (b)-adique de $\mathbb{Z}$ s'appelle *l'anneau des entiers b-adiques*, et est noté $\mathbb{Z}_b$.

On continue en déterminant les éléments inversibles de l'anneau A_I.

1.9. Proposition. *Soit $x = (x_n)_{n\geqslant 0} \in A_I$. Les propriétés suivantes sont équivalentes :*

(1) *on a $x \in A_I^\times$;*

(2) *pour tout $n \geqslant 0$, x est inversible modulo I^{n+1} ;*

(3) *il existe $n \geqslant 0$ tel que x soit inversible modulo I^{n+1} ;*

(4) *la suite x est inversible modulo I.*

Démonstration

(1) $\Longrightarrow$ (2). Soit $x \in A_I^\times$. Alors, il existe un élément $y = (y_n)_{n\geqslant 0} \in A_I$ tel que $xy = ([1]_n)_{n\geqslant 0}$. En particulier, pour tout entier $n \geqslant 0$, on a

$$x_n y_n = [1]_n \in A/I^{n+1},$$

et donc $x_n \in (A/I^{n+1})^\times$. Autrement dit, x est inversible modulo I^{n+1} pour tout $n \geqslant 0$.

(2) $\Longrightarrow$ (3). Évident.

(3) $\Longrightarrow$ (4). Supposons que $x = (x_n)_{n\geqslant 0}$ soit inversible modulo I^{n+1} pour un certain entier $n \geqslant 0$. Par conséquent, il existe une classe $y_n \in A/I^{n+1}$ telle que $x_n y_n = [1]_n$. En utilisant la projection $\pi_{n,0} : A/I^{n+1} \longrightarrow A/I$ et le fait que $\pi_{n,0}(x_n) = x_0$ par cohérence de la suite $(x_n)_{n\geqslant 0}$, on en déduit que $x_0 \in A/I$ est inversible. Ainsi, x est inversible modulo I.

(4) $\Longrightarrow$ (1). Supposons que $x_0 \in (A/I)^\times$. Écrivons $x_n = [a_n]_n$ pour tout entier $n \geqslant 0$. Nous allons montrer l'existence d'une suite $(b_n)_{n\geqslant 0}$ d'éléments de A telle que

(*a*) $b_{n+1} \equiv b_n \mod I^{n+1}$, pour tout $n \geqslant 0$;

(*b*) $a_n b_n \equiv 1 \mod I^{n+1}$, pour tout $n \geqslant 0$.

Par hypothèse, il existe $b_0 \in A$ tel que $a_0 b_0 \equiv 1 \mod I$. Supposons que les éléments $b_0, \ldots, b_n \in A$ soient construits. On cherche $b_{n+1} \in A$ sous la forme

$$b_{n+1} = b_n + \alpha,$$

avec $\alpha \in I^{n+1}$ à déterminer. On a

$$a_{n+1} b_{n+1} = a_{n+1} b_n + a_{n+1}\alpha = (a_{n+1} - a_n) b_n + a_n b_n + a_{n+1}\alpha.$$

Par définition de $b_0, \ldots, b_n$, on a en particulier

$$b_{k+1} \equiv b_k \mod I, \quad \text{pour tout } k \in [\![0, n-1]\!].$$

On en déduit que $b_n \equiv b_0 \mod I$, i.e. $b_n - b_0 \in I$.

Comme la suite $x = ([a_n]_n)_{n\geqslant 0}$ est cohérente, on a $a_{n+1} - a_n \in I^{n+1}$, et donc $(a_{n+1} - a_n)(b_n - b_0) \in I^{n+2}$. On a donc

$$(a_{n+1} - a_n)b_n \equiv (a_{n+1} - a_n)b_0 \mod I^{n+2}.$$

Par cohérence de la suite $x = ([a_n]_n)_{n\geqslant 0}$, on a $a_{n+1} - a_0 \in I$, et par conséquent $(a_{n+1} - a_0)\alpha \in I^{n+2}$, d'où

$$a_{n+1}\alpha \equiv a_0\alpha \mod I^{n+2}.$$

Ainsi, pour tout $\alpha \in I^{n+1}$, on obtient

$$a_{n+1}b_{n+1} - 1 \equiv (a_{n+1} - a_n)b_0 + (a_nb_n - 1) + a_0\alpha \mod I^{n+2}.$$

Posons $\alpha = -b_0\big((a_{n+1} - a_n)b_0 + (a_nb_n - 1)\big)$. Puisque $([a_n]_n)_{n\geqslant 0}$ est cohérente, et puisque $a_nb_n - 1 \in I^{n+1}$ par choix de b_n, on a $\alpha \in I^{n+1}$. Enfin, comme $-a_0b_0 + 1 \in I$ et $(a_{n+1} - a_n)b_0 + (a_nb_n - 1) \in I^{n+1}$, on a

$$a_0\alpha \equiv -\big((a_{n+1} - a_n)b_0 + (a_nb_n - 1)\big) \mod I^{n+2},$$

d'où $a_{n+1}b_{n+1} - 1 \equiv 0 \mod I^{n+2}$. L'élément b_{n+1} vérifie donc les conditions demandées.

L'existence de la suite $(b_n)_{n\geqslant 0}$ possédant les propriétés requises étant établie, posons $y = ([b_n])_{n\geqslant 0}$. Par construction, $y \in A_I$ et est un inverse de x. Cela achève la démonstration. □

1.10. Exemples

(1) Lorsque l'on applique cela à $A[X]$ et $I = (X)$, on retrouve bien, via l'isomorphisme du lemme 1.1, qu'une série formelle est inversible si, et seulement si, son terme constant l'est.

(2) Si $b \geqslant 2$, on obtient qu'un entier b-adique $x = (x_n)_{n\geqslant 0}$ est inversible si, et seulement si, $x_0 \in (\mathbb{Z}/b\mathbb{Z})^\times$.

1.11. Corollaire. *Soit $a \in A$ et soit $n \geqslant 0$ un entier. Alors, a est inversible modulo I^{n+1} si, et seulement si, a est inversible modulo I.*

Démonstration. Le sens direct est clair. Supposons que a soit inversible modulo I. D'après la proposition précédente, $\iota_I(a)$ est inversible. Mais alors, sa n-ième coordonnée $[a]_n$ est un élément inversible de A/I^{n+1}, d'où le résultat. □

1.12. Remarque. Ce corollaire peut bien entendu se démontrer directement à l'aide de calculs similaires à ceux faits dans la démonstration de la proposition 1.9.

2. Propriétés métriques

Dans ce paragraphe, nous allons définir une distance sur A_I et étudier les propriétés de l'espace métrique associé. On commence par une définition.

2.1. Définition. Pour tout $x = (x_n)_{n\geqslant 0} \in A_I$ non nul, on pose

$$v_I(x) = \inf\{n \geqslant 0 \mid x_n \neq [0]_n\}.$$

On pose aussi $v_I(0) = +\infty$.

2.2. Remarque. Pour tout $x \in A_I$, et tout $n \geqslant 0$, on a $v_I(x) \geqslant n$ si, et seulement si, $x_{n-1} = [0]_{n-1}$.

En effet, le sens direct est clair par définition. Inversement, si $x_{n-1} = [0]_{n-1}$, alors on a $x_m = [0]_m$ pour tout $m \in [\![0, n-1]\!]$ par la remarque 1.3. On en déduit alors que $v_I(x) \geqslant n$.

2.3. Exemple. Considérons une série formelle $\sum\limits_{n\geqslant m} a_n X^n \in A[[X]]$ non nulle d'ordre $m \geqslant 0$ (i.e. $a_m \neq 0$). Par le lemme 1.1, elle correspond à l'élément

$$x = \Big(\Big[\sum_{k=m}^{n} a_k X^k\Big]_n\Big)_{n\geqslant 0} \in A[X]_X.$$

Les termes de rang $0, 1, \ldots, m-1$ sont donc nuls. De plus, le terme de rang m, à savoir $\Big[\sum\limits_{k=m}^{m} a_k X^k\Big]_m = [a_m X^m]_m$, est non nul. Ainsi, l'ordre de la série formelle $\sum\limits_{n\geqslant m} a_n X^n \in A[[X]]$, vue comme élément de $A[X]_X$ est égale à m.

2.4. Lemme. *L'application $v_I : A_I \longrightarrow \mathbb{N} \cup \{+\infty\}$ vérifie les propriétés suivantes :*

(1) *pour tout $x \in A_I$, on a $v_I(x) = +\infty$ si, et seulement si, $x = 0$;*

(2) *pour tout $x \in A_I$, on a $v_I(-x) = v_I(x)$;*

(3) *pour tous $x, y \in A_I$, on a $v_I(x+y) \geqslant \min\big(v_I(x), v_I(y)\big)$. De plus, on a égalité si $v_I(x) \neq v_I(y)$;*

(4) *pour tous $x, y \in A_I$, on a[2] $v_I(xy) \geqslant v_I(x) + v_I(y)$.*

De plus, l'application

$$\begin{aligned} d_I : A_I \times A_I &\longrightarrow \mathbb{R}^+ \\ (x, y) &\longmapsto \frac{1}{2^{v_I(x-y)}} \end{aligned}$$

vérifie les propriétés suivantes[3] *:*

2. On utilise la convention $n + (+\infty) = +\infty + n = (+\infty) + (+\infty) = +\infty$.

3. On utilise la convention $\dfrac{1}{2^{+\infty}} = 0$.

(1) *pour tous* $x, y \in A_I$, *on a* $d_I(x,y) = 0$ *si, et seulement si,* $x = y$;
(2) *pour tous* $x, y \in A_I$, *on a* $d_I(y,x) = d_I(x,y)$;
(3) *pour tous* $x, y, z \in A_I$, *on a* $d_I(x,z) \leqslant \max\big(d_I(x,y), d_I(y,z)\big)$.
En particulier, d_I *est une distance sur* A_I.

Démonstration. Commençons par démontrer les propriétés de l'application v_I. Les propriétés (1) et (2) découlent de la définition. Soient $x, y \in A_I$. Les propriétés (3) et (4) étant claires si $x = 0$ ou $y = 0$, on peut supposer que $x = (x_k)_{k\geqslant 0}$ et $y = (y_k)_{k\geqslant 0}$ sont tous deux non nuls.

Posons alors $n = v_I(x)$ et $m = v_I(y)$. Montrons (3). On peut toujours supposer $n \leqslant m$. On a alors

$$x_{n-1} = y_{n-1} = [0]_{n-1},$$

et ainsi $x_{n-1} + y_{n-1} = [0]_{n-1}$. Par la remarque 2.2, on obtient

$$v_I(x+y) \geqslant n.$$

Si maintenant $n < m$, on a $x_n \neq [0]_n$, mais $y_n = [0]_n$, car $n \leqslant m-1$. Ainsi, on a $x_n + y_n = x_n \neq [0]_n$, et par conséquent $v_I(x+y) \leqslant n$, d'où finalement l'égalité $v_I(x+y) = n$ dans ce cas, et le résultat voulu.

Il reste à démontrer (4). Écrivons

$$x_{n+m-1} = [a_{n+m-1}]_{n+m-1} \quad \text{et} \quad y_{n+m-1} = [b_{n+m-1}]_{n+m-1}.$$

Par définition de n et m, on a $x_{n-1} = [0]_{n-1}$ et $y_{m-1} = [0]_{m-1}$. Comme x et y sont des suites cohérentes, on obtient $a_{n+m-1} \in I^n$ et $b_{n+m-1} \in I^m$. Mais alors, $a_{n+m-1}b_{n+m-1} \in I^{n+m}$, d'où $x_{n+m-1}y_{n+m-1} = [0]_{n+m-1}$. Ainsi, $v_I(xy) \geqslant n+m$ par la remarque 2.2.

Les trois premières propriétés de l'application d_I découlent immédiatement de celles de v_I. Le dernier point provient du fait que, pour tous $x, y, z \in A_I$, on a

$$d_I(x,z) \leqslant \max\big(d_I(x,y), d_I(y,z)\big) \leqslant d_I(x,y) + d_I(y,z).$$

Cela achève la démonstration. □

Maintenant que l'on a muni A_I d'une structure d'espace métrique, nous allons pouvoir étudier la convergence des suites d'éléments de A_I. Dans la suite, tous les énoncés relatifs à la notion de convergence le seront pour la topologie induite par la distance d_I.

On commence par caractériser les suites qui convergent vers 0.

2.5. Lemme. *Soit* $(\alpha_k)_{k\geqslant 0}$ *une suite d'éléments de* A_I. *Pour tout* $k \geqslant 0$, *on note* $\alpha_k = (\alpha_{k,n})_{n\geqslant 0}$. *Alors, les propriétés suivantes sont équivalentes :*
(1) *la suite* $(\alpha_k)_{k\geqslant 0}$ *converge vers* 0 *dans* A_I ;

(2) *la suite* $\big(v_I(\alpha_k)\big)_{k\geqslant 0}$ *diverge vers* $+\infty$;

(3) *pour tout* $n \geqslant 0$, *les éléments* $\alpha_{k,n}$, $k \geqslant 0$ *sont presque tous nuls.*

Démonstration. L'équivalence des deux premières propriétés découle de la définition de d_I. Supposons maintenant que $\big(v_I(\alpha_k)\big)_{k\geqslant 0}$ diverge vers $+\infty$. Soit $n \geqslant 0$. Par hypothèse, il existe un entier $N \geqslant 0$ (dépendant de n) tel que pour tout $k \geqslant N$, on ait $v_I(\alpha_k) \geqslant n+1$. En particulier, on a $\alpha_{k,n} = [0]_n$ par la remarque 2.2. Comme cela est vrai pour tout $k \geqslant N$, on obtient bien le résultat souhaité. Inversement, supposons que pour tout $n \geqslant 0$, les éléments $\alpha_{k,n}$, $k \geqslant 0$ soient presque tous nuls. Soit $n \geqslant 0$. Alors, il existe N tel que $\alpha_{k,n} = [0]_n$ pour tout $k \geqslant N$. D'après la remarque 2.2, on a donc $v_I(\alpha_k) \geqslant n+1$ pour tout $k \geqslant N$. Cela démontre que $\big(v_I(\alpha_k)\big)_{k\geqslant 0}$ diverge vers $+\infty$. □

2.6. Remarque. Grâce à la remarque 2.3 et au lemme 2.4, on voit que l'application

$$\begin{array}{rccc} \delta : & A[[X]] \times A[[X]] & \longrightarrow & \mathbb{R} \\ & (S_1, S_2) & \longmapsto & \dfrac{1}{2^{v_X(S_1-S_2)}} \end{array}$$

est une distance sur $A[[X]]$. De plus, l'isomorphisme

$$\varphi : A[[X]] \xrightarrow{\sim} A[X]_X$$

défini dans le lemme 1.1 est en fait une isométrie d'espaces métriques.

L'étude des propriétés métriques de l'anneau $(A[[X]], \delta)$ peuvent alors se traduire au niveau de l'anneau $(A[X]_X, d_X)$ et vice-versa. Par exemple, une suite de séries formelles $(S_k)_{k\geqslant 0}$ converge vers S si, et seulement si, la suite $\big(\varphi(S_k)\big)_{k\geqslant 0}$ converge vers $\varphi(S)$.

De plus, φ étant un morphisme d'anneaux continu, le même résultat est vrai pour les séries : une série $\sum\limits_{k\geqslant 0} S_k$ converge si, et seulement si, $\sum\limits_{k\geqslant 0} \varphi(S_k)$ converge.

Ce point de vue va permettre entre autres d'éclairer sous un nouveau jour la notion de famille sommable, comme nous le verrons plus loin.

On commence par retraduire le lemme précédent dans le cas des séries formelles.

2.7. Exemple. Soit $(S_k)_{k\geqslant 0}$ une suite de séries formelles.

Si $S_k = \sum\limits_{n\geqslant 0} a_{k,n} X^n$, il lui correspond l'élément

$$\alpha_k = \Big(\Big[\sum_{j=0}^{n} a_{k,j} X^j\Big]_n\Big)_{n\geqslant 0}.$$

Alors, $\alpha_{k,n} = \Big[\sum_{j=0}^{n} a_{k,j}X^j\Big]_n$ est nul si, et seulement si, $a_{k,j}$ est nul pour tout entier $j \in [\![0,n]\!]$. On en déduit alors que, pour tout entier $n \geqslant 0$, les éléments $\alpha_{k,n}$, $k \geqslant 0$ sont presque tous nuls si, et seulement si, pour tout entier $n \geqslant 0$, les éléments $a_{k,n}$, $k \geqslant 0$ sont presque tous nuls.

En effet, si pour tout $n \geqslant 0$, $\alpha_{k,n}$ est nul pour tout $k \geqslant k_0(n)$, alors en particulier $a_{k,n}$ est nul pour tout $k \geqslant k_0(n)$ d'après ce qui précède. Inversement, supposons que pour tout $n \geqslant 0$, on ait $a_{k,n} = 0$ pour tout entier $k \geqslant k_0'(n)$. Soit $k \geqslant \max\big(k_0'(0), \ldots, k_0'(n)\big)$. Pour tout $j \in [\![0,n]\!]$, on a alors $a_{k,j} = 0$. Ainsi, $\alpha_{k,n} = 0$.

D'après la remarque 2.6, la suite $(S_k)_{k\geqslant 0}$ converge vers 0 (pour la métrique δ) si, et seulement si, la suite $(\alpha_k)_{k\geqslant 0}$ converge vers 0 dans $A[X]_X$. Au vu de ce qui précède, on en déduit que $(S_k)_{k\geqslant 0}$ converge vers 0 si, et seulement si, la famille $(S_k)_{k\geqslant 0}$ est sommable.

Continuons par un résultat de densité.

2.8. Proposition. *L'image du morphisme d'anneaux $\iota_I : A \longrightarrow A_I$ est dense dans A_I.*

Démonstration. Soit $x = ([a_n]_n)_{n\geqslant 0}$. Pour tout $k \geqslant 0$, posons

$$\alpha_k = \iota_I(a_k) = ([a_k])_{n\geqslant 0} \in \mathrm{Im}(\iota_I).$$

La coordonnée d'indice k de $\alpha_k - x$ est donc $[0]_k$. Par la remarque 2.2, pour tout $k \geqslant 0$, on a ainsi $v(\alpha_k - x) \geqslant k+1$. En particulier, $\big(v_I(\alpha_k - x)\big)_{k\geqslant 0}$ diverge vers $+\infty$, et la suite $(\alpha_k)_{k\geqslant 0}$ est donc une suite d'éléments de $\mathrm{Im}(\iota_I)$ convergeant vers x par le lemme 2.5. Cela achève la démonstration. □

Nous en venons au premier résultat important de cette section.

2.9. Théorème. *L'espace métrique (A_I, d_I) est complet.*

Démonstration. Soit $(\alpha_k)_{k\geqslant 0}$ une suite de Cauchy. Pour tout $n \geqslant 0$, il existe $N_n \geqslant 0$ tel que, pour tous $k \geqslant \ell \geqslant N_n$, on ait $d_I(\alpha_k, \alpha_\ell) \leqslant \frac{1}{2^{n+1}}$, c'est-à-dire $v_I(\alpha_k - \alpha_\ell) \geqslant n+1$.

On définit une suite d'entiers $(N_n')_{n\geqslant 0}$ en posant $N_0' = N_0$ et pour tout entier $n \geqslant 0$, $N_{n+1}' = \max(N_{n+1}, N_n')$. Par définition, $(N_n')_{n\geqslant 0}$ est croissante, et pour tout $n \geqslant 0$, et tous $k \geqslant \ell \geqslant N_n'$, on a $v_I(\alpha_k - \alpha_\ell) \geqslant n+1$.

Écrivons alors $\alpha_k = ([a_{k,n}]_n)_{n\geqslant 0}$, et pour tout $n \geqslant 0$, posons $x_n = [a_{N'_n,n}]_n$. Puisque $N'_{n+1} \geqslant N'_n$, on a donc $v_I(\alpha_{N'_{n+1}} - \alpha_{N'_n}) \geqslant n+1$. En particulier, on a $[a_{N'_{n+1},n} - a_{N'_n,n}]_n = [0]_n$. Mais, $\alpha_{N'_{n+1}}$ étant une suite cohérente, on a aussi $[a_{N'_{n+1},n+1} - a_{N'_{n+1},n}]_n = [0]_n$. En additionnant ces deux égalités, on en déduit donc que $[a_{N'_{n+1},n+1} - a_{N'_n,n}]_n = [0]_n$. Autrement dit, $x = (x_n)_{n\geqslant 0}$ est une suite cohérente.

De plus, par construction, la coordonnée d'indice n de $\alpha_{N'_n} - x$ est nulle, et donc $v_I(\alpha_{N'_n} - x) \geqslant n+1$ par la remarque 2.2. De plus, pour tout entier $k \geqslant N'_n$, on a aussi $v_I(\alpha_k - \alpha_{N'_n}) \geqslant n+1$. Mais alors, pour tout entier $k \geqslant N'_n$, on a

$$v_I(\alpha_k - x) \geqslant \min\left(v_I(\alpha_k - \alpha_{N'_n}), v_I(\alpha_{N'_n} - x)\right) \geqslant n+1.$$

On en déduit donc que $\left(v_I(\alpha_k - x)\right)_{k\geqslant 0}$ diverge vers $+\infty$, i.e. que $(\alpha_k)_{k\geqslant 0}$ converge vers x. Cela achève la démonstration. □

On continue en donnant une condition nécessaire et suffisante pour qu'une suite soit de Cauchy.

2.10. Lemme. *Une suite $(\alpha_k)_{k\geqslant 0}$ d'éléments de A_I est de Cauchy si, et seulement si, la suite $\left(v_I(\alpha_{k+1} - \alpha_k)\right)_{k\geqslant 0}$ diverge vers $+\infty$.*

Démonstration. Soit $(\alpha_k)_{k\geqslant 0}$ une suite de Cauchy. Soit $M \geqslant 0$. Par hypothèse, il existe $N \geqslant 0$ tel que, pour tous n, $m \geqslant N$, on a $d(x_n, x_m) \leqslant \frac{1}{2^M}$. En particulier, pour tout $k \geqslant N$, on a

$$d(x_{k+1}, x_k) \leqslant \frac{1}{2^M},$$

soit $v_I(x_{k+1} - x_k) \geqslant M$. Ainsi, $\left(v_I(\alpha_{k+1} - \alpha_k)\right)_{k\geqslant 0}$ diverge vers $+\infty$.

Inversement, supposons que la suite $\left(v_I(\alpha_{k+1} - \alpha_k)\right)_{k\geqslant 0}$ diverge vers $+\infty$. Soit $\varepsilon > 0$. Soit $M > 0$ tel que $\frac{1}{2^M} < \varepsilon$. Par hypothèse, il existe $N \geqslant 0$ tel que pour tout $k \geqslant N$, on a $v_I(x_{k+1} - x_k) \geqslant M$, et donc

$$d(x_{k+1}, x_k) \leqslant \frac{1}{2^M} < \varepsilon, \quad \text{pour tout } k \geqslant N.$$

Pour tous $n \geqslant m \geqslant N+1$, on a alors

$$d_I(x_n, x_m) \leqslant \max\left(d_I(x_n, x_{n-1}), \ldots, d_I(x_{m+1}, x_m)\right) < \varepsilon,$$

d'où le résultat voulu. □

Le résultat suivant donne une caractérisation surprenante de la convergence d'une série d'éléments de A_I.

2.11. Corollaire. *Soit $(\alpha_k)_{k\geqslant 0}$ une suite d'éléments de A_I.*

Alors, la série $\sum_{k\geqslant 0}\alpha_k$ converge si, et seulement si, la suite $(\alpha_k)_{k\geqslant 0}$ converge vers 0.

Démonstration. Pour tout $k\geqslant 0$, on pose $S_k=\sum_{i=0}^{k}\alpha_i$. Si $(S_k)_{k\geqslant 0}$ converge, alors $(S_k-S_{k-1})_{k\geqslant 1}$ converge vers 0. Autrement dit, $(\alpha_k)_{k\geqslant 1}$ converge vers 0, ou, ce qui revient au même, $(\alpha_k)_{k\geqslant 0}$ converge vers 0. Inversement, supposons que $(\alpha_k)_{k\geqslant 0}$ converge vers 0, c'est-à-dire que $\big(v_I(\alpha_k)\big)_{k\geqslant 0}$ diverge vers $+\infty$ par le lemme 2.5. On a

$$v_I(S_{k+1}-S_k)=v_I(\alpha_{k+1}),$$

et donc $\big(v_I(S_{k+1}-S_k)\big)_{k\geqslant 0}$ diverge vers plus l'infini. Par le lemme 2.10, la suite $(S_k)_{k\geqslant 0}$ est une suite de Cauchy, donc convergente puisque A_I est complet par le théorème 2.9. Autrement dit, $\sum_{k\geqslant 0}\alpha_k$ est une série convergente. Cela achève la démonstration. □

2.12. Exemples

(1) Soit $(S_k)_{k\geqslant 0}$ une suite de séries formelles.

D'après la remarque 2.6 et le corollaire 2.11, la série $\sum_{k\geqslant 0}S_k$ converge (pour la métrique δ sur $A[[X]]$) si, et seulement si, $(S_k)_{k\geqslant 0}$ converge vers 0, i.e. si, et seulement si, la famille $(S_k)_{k\geqslant 0}$ est sommable par l'exemple 2.7.

(2) Soit $\pi\in I$. Alors, pour tout $a\in A$, et tout $n\geqslant 0$, on a

$$v_I\big(a\iota_I(\pi)^n\big)\geqslant nv_I\big(\iota_I(\pi)\big)\geqslant n.$$

Si $S=\sum_{n\geqslant 0}a_nX^n\in A[[X]]$ est une série formelle, le corollaire précédent montre alors que la série $S(\pi)=\sum_{n\geqslant 0}a_n\iota(\pi)^n$ converge dans A_I. On a d'ailleurs

$$S(\pi)=\Big(\Big[\sum_{k=0}^{n}a_k\pi^k\Big]_n\Big)_{n\geqslant 0}.$$

Pour tout $\pi\in I$, on obtient alors un morphisme d'évaluation

$$\begin{aligned} ev_\pi : A[[X]] &\longrightarrow A_I \\ S &\longmapsto S(\pi). \end{aligned}$$

De plus, si S est d'ordre m, alors le terme de rang $m-1$ de $S(\pi)$ est nul, et donc $v_I\big(S(\pi)\big)\geqslant v_X(S)$. Cela entraîne que ev_π est continu,

puisque pour tous S_1, $S_2 \in A[[X]]$, on a alors

$$d_I\big(S_1(\pi), S_2(\pi)\big) \leqslant \delta(S_1, S_2).$$

En particulier, si $(S_k)_{k\geqslant 0}$ est une famille sommable de séries formelles, alors la série $\sum\limits_{k\geqslant 0} S_k(\pi)$ est convergente.

2.13. Remarque. Supposons que $\bigcap\limits_{n\geqslant 0} I^{n+1} = (0)$. Le morphisme d'anneaux $\iota_I : A \longrightarrow A_I$ est alors injectif. Pour tout $a \in A$, posons

$$v'_I(a) = \inf\{n \geqslant 0 \mid a \notin I^{n+1}\}.$$

On pose aussi $v'_I(0) = +\infty$. On vérifie aisément que l'application

$$\begin{aligned} d'_I : A \times A &\longrightarrow \mathbb{R}^+ \\ (a, a') &\longmapsto \frac{1}{2^{v'_I(a-a')}} \end{aligned}$$

est une distance sur A, et que

$$d_I\big(\iota_A(a), \iota_A(a')\big) = d'_I(a, a'), \qquad \text{pour tous } a,\ a' \in A.$$

Autrement dit, $\iota_I : (A, d'_I) \longrightarrow (A_I, d_I)$ est une isométrie dont l'image est dense dans A_I. L'espace métrique (A_I, d_I) étant de plus complet par le théorème 2.9, le triplet (A_I, d_I, ι_I) est donc le complété de l'espace métrique (A, d'_I).

On continue par un lemme.

2.14. Lemme. *Soit $n \geqslant 0$. On munit A/I^{n+1} de la topologie discrète. Alors, le morphisme $\pi_{I,n} : A_I \longrightarrow A/I^{n+1}$ est continu.*

Démonstration. Considérons la métrique triviale $d_0 : A/I^{n+1} \longrightarrow \{0, 1\}$. Autrement dit, pour tous z, $z' \in A/I^{n+1}$, on a

$$d_0(z, z') = \delta_{z,z'}.$$

Alors, d_0 induit la topologie discrète sur A/I^{n+1}. Remarquons maintenant que pour tous x, $y \in A_I$ tels que $\pi_{I,n}(x) \neq \pi_{I,n}(y)$, on a

$$d_I(x, y) \geqslant \frac{1}{2^n},$$

puisqu'en effet, on a alors $v_I(x-y) \leqslant n$ par définition de v_I. En particulier, on a

$$d_0\big(\pi_{I,n}(x), \pi_{I,n}(y)\big) \leqslant 2^n d_I(x, y), \qquad \text{pour tous } x, y \in A_I.$$

La projection $\pi_{I,n}$ est donc 2^n-lipschitzienne, donc a fortiori continue. Cela achève la démonstration. □

On peut maintenant montrer un résultat de compacité.

2.15. Proposition. *L'espace métrique (A_I, d_I) est compact si, et seulement si, A/I^{n+1} est fini pour tout $n \geqslant 0$. Dans ce cas, A_I est totalement discontinu.*

Démonstration. Supposons que A_I soit compact et soit $n \geqslant 0$. D'après le lemme précédent, la projection $\pi_{I,n} : A_I \longrightarrow A/I^{n+1}$ est continue, A/I^{n+1} étant muni de la topologie discrète. Comme la topologie discrète est séparée, l'image de $\pi_{I,n}$ est compacte. Or, la projection $\pi_{I,n}$ est surjective par la remarque 1.7. Ainsi, A/I^{n+1} est compact et discret, donc fini.

Inversement, supposons que A/I^{n+1} soit fini pour tout $n \geqslant 0$. Soit $(\alpha_k)_{k\geqslant 0}$ une suite d'éléments de A_I. On doit montrer que l'on peut en extraire une sous-suite convergente. Puisque A/I est fini, il existe $x_0 \in A/I$ et une partie infinie I_0 de $\mathbb{N}$ telle que $\alpha_{k,0} = x_0$ pour tout $k \in I_0$. De même, puisque l'anneau A/I^2 est fini et I_0 est infini, il existe $x_1 \in A/I^2$ et un sous-ensemble $I_1 \subset I_0$ infini tel que $\alpha_{k,1} = x_1$ pour tout $k \in I_1$. Par récurrence, on construit une suite décroissante $(I_n)_{n\geqslant 0}$ de sous-ensembles infinis de $\mathbb{N}$, et une suite d'éléments $(x_n)_{n\geqslant 0} \in \prod_{n\geqslant 0} A/I^{n+1}$ tels que pour tout $n \geqslant 0$, et tout $k \in I_n$, on ait $\alpha_{k,n} = x_n$. Remarquons que $x = (x_n)_{n\geqslant 0}$ est une suite cohérente. En effet, soit $n \geqslant 0$, et soit $k \in I_{n+1}$. Alors, on a $x_{n+1} = \alpha_{k,n+1}$. Comme $k \in I_{n+1} \subset I_n$, on a alors $x_n = \alpha_{k,n}$. La suite (α_k) étant cohérente, on en déduit que la projection de x_{n+1} sur A/I^{n+1} est x_n. Ainsi, $x \in A_I$.

Pour tout $n \geqslant 0$, on ordonne les éléments de I_n par ordre croissant, et l'on note $\psi(n)$ le n-ième élément de I_n. Montrons alors que la suite extraite $(\alpha_{\psi(k)})_{k\geqslant 0}$ converge vers x. Pour tout $k \geqslant 0$, on a $\psi(k) \in I_k$, et donc $\alpha_{\psi(k),k} = x_k$. Autrement dit, on a

$$v_I(\alpha_{\psi(k)} - x) \geqslant k + 1, \quad \text{pour tout } k \geqslant 0,$$

par la remarque 2.2. Le lemme 2.5 permet alors de conclure.

Montrons la totale discontinuité de A_I dans ce cas. Soit $x = (x_n)_{n\geqslant 0} \in A_I$, et soit C_x sa composante connexe. Pour tout $n \geqslant 0$, le morphisme

$$\pi_{I,n} : A_I \longrightarrow A/I^{n+1}$$

est continu par le lemme 2.14. Ainsi, $\pi_{I,n}(C_x)$ est connexe. Comme l'anneau A/I^{n+1} est fini, $\pi_{I,n}(C_x)$ est réduit à un singleton. Or, $x \in C_x$ et donc $\pi_{I,n}(C_x) = \{\pi_{I,n}(x)\} = \{x_n\}$. Si maintenant $y = (y_n)_{n\geqslant 0} \in C_x$, on en déduit que $y_n = \pi_{I,n}(y) = x_n$, et cela pour tout $n \geqslant 0$. Autrement dit, on a $y = x$, i.e. $C_x = \{x\}$. Cela achève la démonstration. □

2.16. Exemples. Une application de la proposition précédente nous donne les résultats suivants.

(1) Pour tout $b \geqslant 2$, $\mathbb{Z}_b$ est compact et totalement discontinu.

(2) L'anneau $A[[X]]$ est compact si, et seulement si, A est fini.
D'après la remarque 2.6, il faut et il suffit de montrer que $A[X]_X$ est compact si, et seulement si, A est fini. Or, si $A[X]_X$ est compact, l'anneau $A[X]/(X)$, qui est isomorphe à $\simeq A$, est fini. Ainsi, A est fini. Inversement, si A est fini, alors $A[X]/(X^{n+1})$ est fini pour tout $n \geqslant 0$, car toute classe est représentée par un polynôme de degré au plus n. Or, il y a au plus $|A|^{n+1}$ polynômes de degré au plus n à coefficients dans A. Par conséquent, $A[X]_X$ est compact dans ce cas.

On achève ce paragraphe en démontrant un des avatars du lemme de Hensel.

2.17. Théorème (Lemme de Hensel). *Soit $P \in A_I[X]$ un polynôme. On suppose qu'il existe $\alpha_0 \in A_I$ et un entier $m \geqslant 0$ tel que $P'(\alpha_0)$ soit inversible modulo I et $P(\alpha_0) \equiv 0 \mod I^{m+1}$. Alors, il existe un unique élément $\alpha \in A_I$ tel que $P(\alpha) = 0$ et $\alpha \equiv \alpha_0 \mod I^{m+1}$. De plus, α est une racine simple de P.*

Démonstration. Avant de commencer la démonstration proprement dite, remarquons que, pour tout polynôme $Q \in A_I[X]$, et tout $a \in A$, il existe un polynôme $R \in A_I[X]$ tel que

$$Q(X) = Q(a) + Q'(a)(X - a) + (X - a)^2 R.$$

Il suffit en effet de le montrer lorsque $Q = X^k = \big(a + (X - a)\big)^k$, $k \geqslant 0$, auquel cas le résultat est clair en utilisant la formule du binôme de Newton, et de conclure par linéarité. Par évaluation, on obtient donc le résultat suivant : pour tous $x, a \in A_I$ il existe $z \in A_I$ tel que

$$Q(x) = Q(a) + Q'(a)(x - a) + (x - a)^2 z.$$

En particulier, on a $Q(x) - Q(a) = (x - a)y$, $y \in A_I$.

Le point crucial est de démontrer l'existence d'une suite $(x_n)_{n \geqslant 0}$ d'éléments de A_I telle que

(1) $x_0 = \alpha_0$;

(2) pour tout $n \geqslant 0$, on a $P(x_n) \equiv 0 \mod I^{2^n(m+1)}$;

(3) pour tout $n \geqslant 0$, on a $P'(x_n) \in A_I^\times$;

(4) pour tout $n \geqslant 0$, on a $x_{n+1} = x_n - P'(x_n)^{-1} P(x_n)$ [4].

4. Le lecteur pourra faire le rapprochement avec la méthode de Newton pour l'approximation de zéros d'une fonction.

Supposons l'existence d'une telle suite acquise, et déduisons-en le théorème. Remarquons que pour tout $n \geqslant 0$, on a $x_{n+1} - x_n \equiv 0 \mod I^{2^n(m+1)}$ par les points (2) et (3).

On a donc

$$v_I(x_{n+1} - x_n) \geqslant 2^n(m+1) + 1, \quad \text{pour tout } n \geqslant 0.$$

Ainsi, $(x_n)_{n\geqslant 0}$ est une suite de Cauchy par le lemme 2.10, et est donc convergente vers un élément $\alpha \in A_I$ par le théorème 2.9.

De plus, pour tout $n \geqslant 0$, on a $x_{n+1} - x_n \in I^{2^n(m+1)} \subset I^{m+1}$. On en déduit aisément que pour tout $k \geqslant 0$, on a $x_k - x_0 = x_k - \alpha_0 \in I^{m+1}$, soit encore $v_I(x_k - \alpha_0) \geqslant m+2$, ce qui se récrit

$$d_I(x_k, \alpha_0) \leqslant \frac{1}{2^{m+2}}, \quad \text{pour tout } k \geqslant 0.$$

Par passage à la limite (et par continuité d'une distance), on obtient

$$d_I(\alpha, \alpha_0) \leqslant \frac{1}{2^{m+2}},$$

soit $v_I(\alpha, \alpha_0) \geqslant m+2$. Autrement dit, $\alpha \equiv \alpha_0 \mod I^{m+1}$.

Montrons maintenant que $P(\alpha) = 0$. Puisque $P(\alpha) - P(x_n) = (\alpha - x_n)y$, pour un certain $y \in A_I$, on a

$$v_I\big(P(\alpha) - P(x_n)\big) \geqslant v_I(\alpha - x_n), \quad \text{pour tout } n \geqslant 0.$$

Comme $(x_n)_{n\geqslant 0}$ converge vers α, on en déduit que $\big(P(x_n)\big)_{n\geqslant 0}$ converge[5] vers $P(\alpha)$. Mais la condition (2) montre que $v_I\big(P(x_n)\big) \geqslant 2^n(m+1)$ pour tout $n \geqslant 0$, et donc que $\big(P(x_n)\big)_{n\geqslant 0}$ converge vers 0. Par unicité de la limite dans un espace métrique, on obtient $P(\alpha) = 0$.

Il reste à établir l'existence de la suite $(x_n)_{n\geqslant 0}$.

Posons $x_0 = \alpha_0$. Soit $n \geqslant 0$ un entier, et supposons avoir construit des éléments $x_0, \ldots, x_n \in A_I$ vérifiant les conditions (2) et (3). Montrons alors que l'élément $x_{n+1} = x_n - P'(x_n)^{-1}P(x_n)$ vérifie lui aussi ces mêmes conditions. Par (2), on a $x_{k+1} \equiv x_k \mod I^{2^k(m+1)}$ pour tout entier $k \in [\![0, n]\!]$. En particulier, on a $x_{k+1} \equiv x_k \mod I$ pour tout entier $k \in [\![0, n]\!]$, et par suite $x_{n+1} \equiv \alpha_0 \mod I$. On obtient alors $P'(x_{n+1}) \equiv P'(\alpha_0) \mod I$. Comme $P'(\alpha_0)$ est inversible dans A_I, il est inversible modulo I. Par conséquent, $P'(x_{n+1})$ est aussi inversible modulo I, donc inversible dans A_I, et cela d'après la proposition 1.9. Enfin, les considérations faites en début de démonstration, appliquées à $Q = P, x = x_{n+1}$ et $a = x_n$, montrent qu'il existe un élément $z \in A_I$ tel que $P(x_{n+1}) = \big(P'(x_n)P(x_n)\big)^2 z$.

5. Cela provient plus généralement de la continuité de l'addition et du produit dans l'anneau A_I (cf. exercice 2).

Comme $P(x_n) \equiv 0 \mod I^{2^n(m+1)}$ par (2), on a $P(x_n)^2 \equiv 0 \mod I^{2^{n+1}(m+1)}$ (il suffit de prendre un représentant de $\pi_{I,n}(P(x_n))$ pour s'en convaincre). En particulier, on a

$$P(x_{n+1}) \equiv 0 \mod I^{2^{n+1}(m+1)}.$$

Cela démontre par récurrence l'existence de la suite $(x_n)_{n\geqslant 0}$ ayant les propriétés voulues, d'où l'existence de α. Si α était une racine multiple de P, ou s'il existait $\beta \in A_I$ distinct de α vérifiant les même propriétés, alors on aurait $P = (X-\alpha)^2 Q$ ou $P = (X-\alpha)(X-\beta)Q$, $Q \in A_I[X]$. Or, α et β sont congrus à α_0 modulo I^{m+1}. Par réduction modulo I^{m+1}, on obtiendrait que α_0 serait une racine multiple de P modulo I^{m+1}, donc aussi modulo I, et cela dans les deux cas. En particulier, $P'(\alpha_0)$ serait nul modulo I, d'où une contradiction. Cela achève la démonstration. □

3. Complétion en un idéal principal

Nous allons maintenant étudier le cas où I est un idéal principal. Notons que si $I = (\pi)$, on a $I^n = (\pi^n)$ pour tout $n \geqslant 0$.

L'anneau A_I est alors noté A_π, et les applications v_I et d_I sont respectivement notées v_π et d_π.

Nous allons retraduire les résultats des paragraphes précédents dans ce cas, et même les améliorer un peu, dans le cas où π est régulier dans le sens suivant.

3.1. Définition. On dit que $\pi \in A$ est *régulier* si, pour tous a, $b \in A$, on a

$$\pi a = \pi b \quad \Longrightarrow \quad a = b.$$

Autrement dit, π est régulier si l'application

$$\begin{aligned} \ell_\pi : A &\longrightarrow A \\ a &\longmapsto \pi a \end{aligned}$$

est injective.

3.2. Exemples

(1) Pour tout anneau A, l'élément 1_A régulier.

(2) Un diviseur de zéro d'un anneau non trivial n'est jamais régulier.

(3) Si A est intègre, tout élément non nul est régulier.

(4) Pour tout anneau A, l'élément $X \in A[X]$ est régulier, de même que tout polynôme de $A[X]$ dont le coefficient dominant n'est pas un diviseur de zéro, d'après le lemme XV-6.8.

(5) Si π_1, $\pi_2 \in A$ sont réguliers, alors $\pi_1\pi_2 \in A$ est régulier. En particulier, si $\pi \in A$ est régulier, l'élément π^n est régulier pour tout $n \geqslant 0$.

On a alors le théorème suivant.

3.3. Théorème. *Soit $\pi \in A$. L'application $v_\pi : A_\pi \longrightarrow \mathbb{N} \cup \{+\infty\}$ vérifie les propriétés suivantes :*

(1) *pour tout $x \in A_\pi$, on a $v_\pi(x) = +\infty$ si, et seulement si, $x = 0$;*

(2) *pour tout $x \in A_\pi$, on a $v_\pi(-x) = v_\pi(x)$;*

(3) *pour tous $x, y \in A_\pi$, on a $v_\pi(x+y) \geqslant \min\big(v_\pi(x), v_\pi(y)\big)$. De plus, on a égalité si $v_\pi(x) \neq v_\pi(y)$;*

(4) *pour tous $x, y \in A_\pi$, $v_\pi(xy) \geqslant v_\pi(x) + v_\pi(y)$.*
Si de plus $\pi \in A$ est un élément premier régulier, alors pour tous éléments $x, y \in A_\pi$, on a[6] *$v_\pi(xy) = v_\pi(x) + v_\pi(y)$.*

De plus :

(i) *on a $x \in A_\pi^\times$ si, et seulement si, x est inversible modulo π ;*

(ii) *si π est un élément premier régulier, l'anneau A_π est intègre ;*

(iii) *l'espace métrique (A_π, d_π) est complet ;*

(iv) *l'image du morphisme $\iota_\pi : A \longrightarrow A_\pi$ est dense dans A_π. De plus, ι_π est injectif si, et seulement si, $\bigcap_{n\geqslant 0} (\pi^{n+1}) = (0)$. En particulier, c'est le cas si A est factoriel ;*

(v) *l'anneau A_π est compact si, et seulement si, $A/(\pi)$ est fini. Dans ce cas, A_π est totalement discontinu.*

Démonstration. Les points (1)-(3) découlent du lemme 2.4. Par ailleurs, d'après ce même lemme, on sait déjà que pour tous $x, y \in A_\pi$, on a

$$v_\pi(xy) \geqslant v_\pi(x) + v_\pi(y),$$

d'où la première partie de (4).

Supposons maintenant que π soit un élément premier régulier. Pour démontrer qu'il y égalité dans ce cas, on peut toujours supposer que x et y sont non nuls, sinon le résultat est clair. Écrivons $x = ([a_k]_k)_{k\geqslant 0}$ et $y = ([b_k]_k)_{k\geqslant 0}$. Soit $n = v_\pi(x)$ et $m = v_\pi(y)$. Puisque $n = v_\pi(x)$, on a $a_n \notin (\pi^{n+1})$. Par cohérence, on a $a_n \equiv a_{n-1} \equiv 0\ [\pi^n]$, et donc $a_n \in (\pi^n)$.

On en déduit $a_n = u\pi^n$, où $u \in A$ vérifie $\pi \nmid u$. En effet, si l'on avait $\pi \mid u$, on aurait $a_n \in (\pi^{n+1})$. Par cohérence, on a $a_{n+m} \equiv a_n\ [\pi^{n+1}]$. Il existe donc $u' \in A$ tel que

$$a_{n+m} = \pi^n u + \pi^{n+1}u' = \pi^n(u + \pi u').$$

Autrement dit, on peut écrire $a_{n+m} = \pi^n a'$, avec $\pi \nmid a'$. De même, on peut écrire $b_{n+m} = \pi^m b'$, avec $\pi \nmid b'$. Mais alors, on a

$$a_{n+m}b_{n+m} = \pi^{n+m}a'b'.$$

Comme (π) est un idéal premier, on en déduit que $a'b' \notin (\pi)$.

6. L'application v_π est alors ce que l'on appelle une valuation sur A_π.

Par suite, on a $a_{n+m}b_{n+m} \notin (\pi^{n+m+1})$. En effet, supposons au contraire qu'il existe un élément $v \in A$ tel que

$$a_{n+m}b_{n+m} = \pi^{n+m}a'b' = v\pi^{n+m+1}.$$

Puisque π est régulier, π^{n+m} l'est aussi, et l'on obtient $a'b' = v\pi \in (\pi)$, d'où une contradiction. Autrement dit, on a

$$[a_{n+m}]_{n+m}[b_{n+m}]_{n+m} \neq [0]_{n+m}.$$

Par conséquent, on a $v_\pi(xy) \leqslant n+m$, ce qui démontre l'inégalité manquante, d'où (4).

La propriété (i) n'est rien d'autre que la proposition 1.9. Montrons maintenant que l'anneau A_π est intègre lorsque π est un élément premier régulier. Bien entendu, A_π est non trivial, et commutatif. Soient $x, y \in A \setminus \{0\}$. Alors, $v_\pi(x) < +\infty$ et $v_\pi(y) < +\infty$. Par conséquent, on a

$$v_\pi(xy) = v_\pi(x) + v_\pi(y) < +\infty$$

par (4), et ainsi $xy \in A \setminus \{0\}$, d'où l'intégrité de A_π.

De plus, (A_π, d_π) est complet par le théorème 2.9, et le morphisme d'anneaux ι_π est d'image dense par la proposition 2.8. Rappelons maintenant que l'on a $\mathrm{Ker}(\iota_\pi) = \bigcap_{n\geqslant 0} (\pi^{n+1})$. Montrons que cette intersection est nulle lorsque A est factoriel. Soit $a \in \bigcap_{n\geqslant 0} (\pi^{n+1})$. Supposons que a soit non nul. Par factorialité de A, on peut écrire $a = u\pi^n$, avec $n \geqslant 0$ et $\pi \nmid u$. En particulier, on a $a \notin (\pi^{n+1})$, d'où une contradiction.

Il reste à montrer (v). Nous allons vérifier que si l'anneau $A/(\pi)$ est fini, alors $A/(\pi^{n+1})$ est fini pour tout $n \geqslant 0$. La proposition 2.15 nous permettra alors de conclure. Soit $n \geqslant 0$. Le morphisme de groupes

$$\begin{aligned} A &\longrightarrow (\pi^n)/(\pi^{n+1}) \\ a &\longmapsto [\pi^n]_{n+1}[a]_{n+1} \end{aligned}$$

est surjectif, et son noyau contient [7] (π). Le théorème de factorisation induit alors un morphisme de groupes surjectif

$$A/(\pi) \longrightarrow (\pi^n)/(\pi^{n+1}),$$

et l'on a donc $|(\pi^n)/(\pi^{n+1})| \leqslant |A/(\pi)|$ pour tout $n \geqslant 0$. Par ailleurs, on a un isomorphisme

$$\big(A/(\pi^{n+1})\big)/\big((\pi^n)/(\pi^{n+1})\big) \simeq A/(\pi^n).$$

7. Il est même égal à (π) si π est régulier.

Pour tout $n \geqslant 0$, on a ainsi

$$|A/(\pi^{n+1})| = |A/(\pi^n)||(\pi^n)/(\pi^{n+1})| \leqslant |A/(\pi^n)||A/(\pi)|,$$

d'où, finalement,

$$|A/(\pi^{n+1})| \leqslant |A/(\pi)|^n < +\infty, \quad \text{pour tout } n \geqslant 0$$

par une récurrence immédiate. On a ainsi le résultat voulu, ce qui achève la démonstration. □

3.4. Exemples

(1) Si A est un anneau intègre, X est un élément premier régulier de $A[X]$, puisque l'évaluation en 0 induit un isomorphisme $A[X]/(X) \simeq A$. De plus, tout polynôme $P \in A[X]$ non nul s'écrit de manière unique sous la forme $P = QX^m$, avec $X \nmid Q$, $m \geqslant 0$. On en déduit alors aisément que $\bigcap_{n\geqslant 0} (X^{n+1}) = (0)$.

Les résultats précédents et la remarque 2.6 montrent alors que l'anneau $A[[X]]$ est intègre, complet, et que le morphisme d'anneaux

$$\iota_X : A[X] \longrightarrow A[[X]]$$

est injectif d'image dense. De plus, $A[[X]]$ est compact si, et seulement si, A est fini, auquel cas $A[[X]]$ est totalement discontinu. Enfin, une série formelle $S \in A[[X]]$ est inversible si, et seulement si, son terme constant est inversible dans A.

(2) Pour tout nombre premier p, l'anneau $\mathbb{Z}_p$ est intègre, compact, complet, totalement discontinu et l'application $\iota_p : \mathbb{Z} \longrightarrow \mathbb{Z}_p$ est injective, d'image dense. De plus, un élément $x \in \mathbb{Z}_p$ est inversible si, et seulement si, x est non nul modulo p.

Notation. Pour tout $a \in A$, et tout $x \in A_\pi$, on pose $a \cdot x = \iota_\pi(a)x$.

3.5. Lemme. *Soit $(s_n)_{n\geqslant 0}$ une suite d'éléments de A.*

Alors, la série $\sum_{n\geqslant 0} s_n\pi^n \cdot 1_{A_\pi}$ converge vers $\left(\left[\sum_{k=0}^{n} s_k\pi^k\right]_n\right)_{n\geqslant 0}$.

Démonstration. Pour tout entier $n \geqslant 0$, on a $v_\pi(s_n\pi^n \cdot 1_{A_\pi}) \geqslant n$, et donc la suite $(s_n\pi^n \cdot 1_{A_\pi})_{n\geqslant 0}$ converge vers 0. D'après le corollaire 2.11, la série

$$\sum_{n\geqslant 0} s_n\pi^n \cdot 1_{A_\pi}$$

est convergente.

D'après l'exemple 1.4 (2), la suite

$$x = \Big(\Big[\sum_{k=0}^{n} s_k \pi^k \Big]_n \Big)_{n \geqslant 0}$$

est cohérente. On vérifie que la coordonnée d'indice n de $x - \sum_{k=0}^{n} s_k \pi^k \cdot 1_{A_\pi}$ est nulle, donc

$$v_\pi \Big(x - \sum_{k=0}^{n} s_k \pi^k \cdot 1_{A_\pi} \Big) \geqslant n+1, \qquad \text{pour tout } n \geqslant 0.$$

Ainsi, la suite $\Big(\sum_{k=0}^{n} s_k \pi^k \cdot 1_{A_\pi} \Big)_{n \geqslant 0}$ converge vers x, i.e.

$$\sum_{n \geqslant 0} s_n \pi^n \cdot 1_{A_\pi} = x,$$

d'où le résultat. □

On peut maintenant énoncer le résultat suivant.

3.6. Théorème. *Soit $\pi \in A$, et soit $S \subset A$ un système complet de représentants*[8] *de l'anneau quotient $A/(\pi)$.*

Alors, pour tout $x \in A_I$, il existe une suite $(s_n)_{n \geqslant 0}$ d'éléments de S telle que $x = \sum_{n \geqslant 0} s_n \pi^n \cdot 1_{A_\pi}$. Si de plus π est régulier, la suite $(s_n)_{n \geqslant 0}$ est unique.

Démonstration. Soit $x \in A_\pi$, et écrivons $x = ([a_n]_n)_{n \geqslant 0}$. Nous allons construire une suite $(s_n)_{n \geqslant 0}$ d'éléments de S telle que

$$a_n - \sum_{k=0}^{n} s_k \pi^k \in (\pi^{n+1}), \qquad \text{pour tout } n \geqslant 0.$$

Soit $s_0 \in S$ l'unique élément de S tel que $a_0 \equiv s_0 \ [\pi]$. Soit $n \geqslant 0$ un entier. Supposons que les éléments $s_0, \ldots, s_n \in S$ soient construits. Par cohérence, on a $a_{n+1} \equiv a_n \ [\pi^{n+1}]$. Par choix de $s_0, \ldots, s_n$, on a donc

$$a_{n+1} \equiv \sum_{k=0}^{n} s_k \pi^k \ [\pi^{n+1}].$$

Ainsi, il existe $u \in A$ tel que

$$a_{n+1} = \sum_{k=0}^{n} s_k \pi^k + \pi^{n+1} u.$$

Soit $s_{n+1} \in S$ l'unique élément de S tel que $u \equiv s_{n+1} \ [\pi]$. Il existe donc un

8. Autrement dit, un sous-ensemble S de A tel que tout élément $a \in A$ est congru modulo π à un et un seul élément de S.

élément $v \in A$ tel que $u = s_{n+1} + \pi v$, et par conséquent

$$a_{n+1} = \sum_{k=0}^{n} s_k \pi^k + s_{n+1}\pi^{n+1} + \pi^{n+2} v.$$

On a donc bien $a_{n+1} - \sum_{k=0}^{n+1} s_k \pi^k \in (\pi^{n+2})$.

Par récurrence, on obtient donc l'existence la suite $(s_n)_{n\geqslant 0}$ vérifiant les conditions requises. Par définition de cette suite, on a

$$x = \left(\left[\sum_{k=0}^{n} s_k \pi^k \right]_n \right)_{n\geqslant 0}.$$

D'après le lemme précédent, on obtient $x = \sum_{n\geqslant 0} s_n \pi^n \cdot 1_{A_\pi}$.

Supposons que π soit régulier, et soient $(s_n)_{n\geqslant 0}$ et $(s'_n)_{n\geqslant 0}$ deux suites d'éléments de S telles que

$$x = \sum_{n\geqslant 0} s_n \pi^n \cdot 1_{A_\pi} = \sum_{n\geqslant 0} s'_n \pi^n \cdot 1_{A_\pi}.$$

Toujours d'après le lemme précédent, on obtient donc

$$\sum_{k=0}^{n} s_k \pi^k \equiv \sum_{k=0}^{n} s'_k \pi^k \quad [\pi^{n+1}], \qquad \text{pour tout } n \geqslant 0. \qquad (*)$$

Montrons par récurrence que $s_n = s'_n$ pour tout $n \geqslant 0$.

Soit (H_n) la proposition suivante :

(H_n) Pour tout $k \in [\![0, n]\!]$, on a $s_k = s'_k$.

Comme $s_0 \equiv s'_0\ [\pi]$ par $(*)$, on a $s_0 = s'_0$ par choix de S. Ainsi, (H_0) est vraie. Soit $n \geqslant 0$, et supposons que (H_n) soit vraie. L'hypothèse (H_n) et la congruence $(*)$ impliquent la congruence

$$s_{n+1}\pi^{n+1} \equiv s'_{n+1}\pi^{n+1} \quad [\pi^{n+2}].$$

Il existe donc $u \in A$ tel que

$$s_{n+1}\pi^{n+1} = s'_{n+1}\pi^{n+1} + \pi^{n+2} u.$$

Par régularité de π (et donc de π^{n+1}), on obtient $s_{n+1} = s'_{n+1} + \pi u$, soit encore $s_{n+1} \equiv s'_{n+1}\ [\pi]$. Par choix de S, on en déduit que $s_{n+1} = s'_{n+1}$. Ainsi, la propriété (H_{n+1}) est vraie, et cela achève la récurrence, ainsi que la démonstration. □

3.7. Définition. Soit $\pi \in A$ un élément régulier, et soit S un système complet de représentants de $A/(\pi)$. Soit $x \in A_\pi$; écrivons $x = \sum_{n\geqslant 0} s_n\pi^n\cdot 1_{A_\pi}$, avec $s_n \in S$ pour tout $n \geqslant 0$. La série $\sum_{n\geqslant 0} s_n\pi^n\cdot 1_{A_\pi}$ s'appelle le *développement de Hensel* de x (par rapport à S).

3.8. Remarques. Soit $\pi \in A$ un élément régulier, et soit $x = \sum_{n\geqslant 0} s_n\pi^n\cdot 1_{A_\pi}$ le développement de Hensel de $x \in A_\pi$ par rapport à S.

(1) Le lemme 3.5 montre que
$$x = \Big(\Big[\sum_{k=0}^{n} s_k\pi^k\Big]_n\Big)_{n\geqslant 0}.$$
Autrement dit, pour tout $n \geqslant 0$, on a $\pi_{I,n}\Big(\sum_{k\geqslant 0} s_k\pi^k\Big) = \Big[\sum_{k=0}^{n} s_k\pi^k\Big]_n$.

(2) Supposons que $0 \in S$. Si $x \in A_\pi$ est non nul, alors $v_I(x)$ est le plus petit entier $n \geqslant 0$ tel que $s_n \neq 0$.
En effet, si $x \neq 0$, il existe au moins un coefficient $s_n \in S$ non nul. Choisissons n minimal pour cette propriété. Alors,
$$x = \pi^n s_n + \pi^{n+1}\sum_{k\geqslant 0} s_{k+n+1}\pi^k.$$
On a donc $\pi_{I,n-1}(x) = [0]_n \in A/(\pi^n)$, et ainsi $v_\pi(x) \geqslant n$. Par ailleurs, puisque $s_n \neq 0$, par choix de S, cela signifie que $[s_n]_0 \neq [0]_0 \in A/(\pi)$. Autrement dit, $s_n \notin (\pi)$. Mais alors, on a $\pi^n s_n \notin (\pi^{n+1})$ (sinon, on aurait $s \in (\pi)$ par régularité de π^n). Ainsi, on a
$$\pi_{\pi,n}(x) = [\pi^n s_n]_n \neq [0]_n,$$
d'où $v_\pi(x) = n$.

(3) Lorsque $A = \mathbb{Z}$ et $\pi = p$, le développement de Hensel est calculé par rapport à $S = [\![0, p-1]\!]$. Tout élément de $\mathbb{Z}_p$ s'écrit donc de manière unique sous la forme
$$x = \sum_{n\geqslant 0} a_n p^n\cdot 1_{\mathbb{Z}_p}, \quad a_n \in [\![0, p-1]\!].$$

3.9. Exemple. Reprenons le cas de $A[X]_X$. Alors, $A[X]/(X) \simeq A$. Si l'on choisit A comme système complet de représentants de $A[X]/(X)$, on obtient que tout élément $x \in A[X]_X$ peut s'écrire de manière unique sous la forme
$$x = \sum_{n\geqslant 0} a_n X^n\cdot 1_{A[X]_X}.$$
On retrouve que x peut se voir comme une série formelle (quitte à faire un petit abus de notation).

D'ailleurs, si $S = \sum_{n\geqslant 0} a_n X^n$ est une série formelle, alors le développement de Hensel de $\varphi(S) = \Big(\Big[\sum_{k=0}^{n} a_k X^k\Big]_n\Big)_{n\geqslant 0}$ n'est rien d'autre que $\sum_{n\geqslant 0} a_n X^n \cdot 1_{A[X]_X}$.

3.10. Remarque. Soit $\pi \in A$ un élément régulier, et soit S un système complet de représentants de $A/(\pi)$. Supposons que $0 \in S$.

Un élément $x \in A_\pi$ avec un développement de Hensel fini (i.e. tel $s_n = 0$ pour tout n suffisamment grand) est un élément de A (ou plus précisément de $\iota_\pi(A)$). Cette observation facile est laissée au lecteur. En revanche, la réciproque est fausse. Si $A = \mathbb{Z}$, le lecteur vérifiera aisément que le développement de Hensel de $-1 \in \mathbb{Z}_p$ est donné par

$$-1 = \sum_{n\geqslant 0} (p-1)p^n \cdot 1_{\mathbb{Z}_p}.$$

Nous allons finir ce chapitre en étudiant de plus près le cas où $A/(\pi)$ est un corps.

3.11. Proposition. *Soit $\pi \in A$ un élément premier régulier. On suppose de plus que $A/(\pi)$ est un corps. Alors :*

(1) *un élément $x \in A_\pi$ est inversible si, et seulement si, $v_\pi(x) = 0$;*

(2) *tout élément $x \in A_\pi \setminus \{0\}$ s'écrit de manière unique sous la forme*

$$x = \pi^m \cdot u, \quad m \geqslant 0,\ u \in A_\pi^\times.$$

De plus, on a $m = v_\pi(x)$;

(3) *tout idéal non nul de A_π est de la forme $(\pi^n \cdot 1_{A_\pi})$ pour un unique $n \geqslant 0$. De plus, $(\pi \cdot 1_{A_\pi})$ est l'unique idéal maximal de A_π. En particulier, A_π est un anneau principal.*

Démonstration. Le premier point découle du théorème 3.3 (i), et du fait que $A/(\pi)$ est un corps. Soit $x \in A_\pi \setminus \{0\}$. Choisissons un système complet de représentants S de $A/(\pi)$ contenant 0, et écrivons

$$x = \sum_{n\geqslant 0} s_n \pi^n \cdot 1_{A_\pi}.$$

Par la remarque 3.8 (1), on a donc

$$x = \Big(\sum_{k=0}^{n} s_k \pi^k\Big)_{n\geqslant 0}.$$

Soit $m = v_\pi(x)$. Alors, par la remarque 3.8 (2), $s_m \neq 0$ et $s_n = 0$, pour

tout $k \in [\![0, m-1]\!]$. On a ainsi

$$x = \Big(\sum_{k=m}^{n} s_k \pi^k \Big)_{n \geqslant 0}.$$

Posons $u = \Big(\sum_{k=0}^{n} s_{k+m} \pi^k \Big)_{n \geqslant 0}$. On vérifie que $u \in A_\pi$. De plus, on a

$$\pi^k \cdot u = \Big(\sum_{k=0}^{n} s_{k+m} \pi^{k+m} \Big)_{n \geqslant 0} = \Big(\sum_{k=m}^{n} s_k \pi^k \Big)_{n \geqslant 0} = x.$$

Enfin, puisque $s_m \neq 0$, par choix de S, s_m est non nul modulo π, et u est donc inversible par le premier point. Cela démontre l'existence de la décomposition annoncée. Supposons maintenant que l'on ait deux décompositions

$$x = \pi^m \cdot u = \pi^n \cdot v, \quad n, m \geqslant 0, \ u, v \in A_\pi^\times.$$

Remarquons que l'on a $x = (\pi^m \cdot 1_{A_\pi}) u$, et d'après le théorème 3.3 (4), il vient donc

$$v_\pi(x) = v_\pi(\pi^m \cdot 1_{A_\pi}) + v_\pi(u) = v_\pi(\pi^m \cdot 1_{A_\pi}).$$

Or, on a $\pi^m \in (\pi^{k+1})$ pour tout $k \in [\![0, m-1]\!]$ et $\pi^m \notin (\pi^{m+1})$ (par régularité de π^m), et l'on a donc

$$v_\pi(\pi^m \cdot 1_{A_\pi}) = v_\pi\big(([\pi^m]_k)_{k \geqslant 0}\big) = m.$$

Ainsi, on a $v_\pi(\pi^m \cdot u) = m$, et de même $v_\pi(\pi^n \cdot v) = n$. Par conséquent, on obtient $m = n$. Mais alors, on a

$$x = (\pi^m \cdot 1_{A_\pi}) u = (\pi^m \cdot 1_{A_\pi}) v,$$

d'où $u = v$ par intégrité de A_π (cf. théorème 3.3 (ii)). Cela démontre l'unicité de la décomposition, d'où le point (2).

Soit I un idéal non nul de A_π, et soit $x_0 \in I \setminus \{0\}$ tel que $v_\pi(x_0)$ soit minimale. Posons $n = v_\pi(x_0)$, et montrons que $I = (\pi^n \cdot 1_{A_\pi})$. Soit $x \in I$. Si $x = 0$, on a bien entendu $x \in (\pi^n \cdot 1_{A_\pi})$. Si x est non nul, d'après le point précédent, on a

$$x = \pi^m \cdot u = (\pi^m \cdot 1_{A_\pi}) u,$$

avec $u \in A_\pi^\times$ et $m = v_\pi(x)$. Par définition de x_0, on a donc $m \geqslant n$, et ainsi

$$\pi^n \cdot u = (\pi^{m-n} \cdot u)(\pi^n \cdot 1_{A_\pi}) \in (\pi^n \cdot 1_{A_\pi}).$$

Ainsi, $I \subset (\pi^n \cdot 1_{A_\pi})$. Par ailleurs, on a

$$\pi^n \cdot 1_{A_\pi} = (\pi^n \cdot 1_{A_\pi}) u u^{-1} = x_0 u^{-1} \in I,$$

d'où l'inclusion $(\pi^n \cdot 1_{A_\pi}) \subset I$.

Ainsi, on a bien $(\pi^n \cdot 1_{A_\pi}) = I$. Supposons maintenant que l'on ait une

égalité
$$(\pi^n \cdot 1_{A_\pi}) = (\pi^m \cdot 1_{A_\pi}), \quad n, m \geqslant 0.$$
Alors, il existe $x \in A_\pi$ tel que $\pi^n \cdot 1_{A_\pi} = \pi^m \cdot 1_{A_\pi} x$. On a alors
$$n = v_\pi(\pi^n \cdot 1_{A_\pi}) = v_\pi(\pi^m \cdot 1_{A_\pi}) + v_\pi(x) = m + v_\pi(x) \geqslant m.$$
En échangeant les rôles de m et n, on obtient aussi $m \geqslant n$, d'où $n = m$.
Cela montre en particulier que tout idéal distinct de A_π est contenu dans l'idéal $(\pi \cdot 1_{A_\pi})$. Comme cet idéal est aussi distinct de A_π (car $\pi \cdot 1_{A_\pi}$ n'est pas inversible), on en déduit que $(\pi \cdot 1_{A_\pi})$ est un idéal maximal. De plus, si $\mathfrak{m}$ est un idéal maximal, il est contenu dans $(\pi \cdot 1_{A_\pi})$, et par maximalité de $\mathfrak{m}$, on a $\mathfrak{m} = (\pi \cdot 1_{A_\pi})$. Cela achève la démonstration. □

3.12. Exemple. Ce théorème s'applique en particulier à l'anneau $K[[X]]$, où K est un corps, ou à l'anneau $\mathbb{Z}_p$ pour tout p premier.

Nous terminons ce chapitre en faisant quelques remarques sur le corps des fractions de A_π lorsque les hypothèses de la proposition précédente sont vérifiées.

3.13. Remarque. Soit $\pi \in A$ un élément premier régulier. On suppose de plus que $A/(\pi)$ est un corps. On note K_π le corps des fractions de A_π. Alors, l'application
$$\begin{aligned} \tilde{v}_\pi : K_\pi &\longrightarrow \mathbb{Z} \cup \{+\infty\} \\ \frac{a}{b} &\longmapsto v_\pi(a) - v_\pi(b) \end{aligned}$$
est bien définie, et prolonge l'application v_π. On définit alors comme précédemment une application $\tilde{d}_\pi : K_\pi \times K_\pi \longrightarrow \mathbb{R}$ qui prolonge d_π. On vérifie aisément que $\tilde{v}_\pi$ et $\tilde{d}_\pi$ jouissent de propriétés similaires à celles de v_π et d_π respectivement. En particulier, $\tilde{d}_\pi$ est une distance sur K_π. On peut alors démontrer les résultats suivants :

(1) pour tout $x \in K_\pi$, on a $x \in A_\pi$ si, et seulement si, $\tilde{v}_\pi(x) \geqslant 0$;
(2) tout élément $x \in K_\pi \setminus \{0\}$ s'écrit de manière unique sous la forme
$$x = u\pi^m, \quad u \in A_\pi^\times, \ m \in \mathbb{Z}.$$
De plus, $m = \tilde{v}_\pi(x)$;
(3) l'espace métrique $(K_\pi, \tilde{d}_\pi)$ est complet ;
(4) si $\bigcap_{n \geqslant 0} (\pi^{n+1}) = (0)$, et si A est intègre, le morphisme $\iota_\pi : A \longrightarrow A_\pi$ se prolonge de manière unique en un morphisme $\tilde{\iota}_\pi : K \longrightarrow K_\pi$, d'image dense ;
(5) pour toute suite $(x_n)_{n \geqslant 0}$ d'éléments de K_π, la série $\sum_{n \geqslant 0} x_n$ converge si, et seulement si, $(x_n)_{n \geqslant 0}$ converge vers 0 ;

(6) soit S un système complet de représentants de $A/(\pi)$ contenant 0. Alors, tout $x \in K_\pi$ s'écrit de manière unique sous la forme

$$x = \sum_{n \geqslant m} s_n \pi^n \cdot 1_{K_\pi}, \quad m \in \mathbb{Z},\ s_n \in S,\ s_m \neq 0.$$

De plus, $m = \tilde{v}_\pi(x)$.

Les détails sont laissés au lecteur.

3.14. Exemple. Si p est un nombre premier, le corps des fractions de $\mathbb{Z}_p$, noté $\mathbb{Q}_p$, est appelé le *corps des nombres p-adiques*, et est d'une grande importance en théorie des nombres.

4. Exercices

E1. Soit A un anneau commutatif, et soit I un idéal de A. On appelle topologie I-adique sur A l'unique topologie dont une base de voisinages de $a \in A$ est donnée par $(a + I^{n+1})_{n \geqslant 0}$.

a. Montrer que la topologie I-adique est la topologie la moins fine rendant toutes les applications $A \longrightarrow A/I^{n+1}$ continues (lorsque l'on munit l'anneau A/I^{n+1} de la topologie discrète).

b. Montrer que l'addition et le produit sur A sont continues.

c. Montrer que la topologie I-adique est séparée si, et seulement si, on a

$$\bigcap_{n \geqslant 0} I^{n+1} = (0).$$

Montrer que dans ce cas, la topologie I-adique est induite par la distance d'_I définie dans la remarque 2.13.

E2. Soit A un anneau commutatif, et soit I un idéal de A.

Montrer que la somme et le produit dans A_I sont continues, et que la topologie sur A_I est la topologie induite par la topologie produit sur l'anneau $\prod_{n \geqslant 0} A/I^{n+1}$ (obtenue lorsque chaque A/I^{n+1} est munie de la topologie discrète).

E3. Soit A un anneau commutatif, et soit I un idéal de A. On suppose qu'il existe $m \geqslant 0$ tel que $I^{m+1} = (0)$. Montrer que l'on a un isomorphisme d'anneaux $A_I \simeq A$, et que cet isomorphisme est un homéomorphisme.

E4. Soit A un anneau commutatif. Soient I, J deux idéaux comaximaux de A. Montrer que l'application

$$\begin{array}{ccc} A_{IJ} & \longrightarrow & A_I \times A_J \\ ([x_n]_{IJ,n})_{n\geqslant 0} & \longmapsto & (([x_n]_{I,n})_{n\geqslant 0}, ([x_n]_{J,n})_{n\geqslant 0}) \end{array}$$

est un isomorphisme d'anneaux, et que cet isomorphisme est un homéomorphisme.

E5. Soit $b \geqslant 2$. Montrer que $\mathbb{Z}_b$ est intègre si, et seulement si, b est une puissance d'un nombre premier.

E6. Soient $k, m \geqslant 0$ deux entiers. On dit que k se termine par m si le développement décimal de k se termine par le développement décimal de m. Par exemple, 128467 se termine par 467 (ainsi que par 67, et par 7).

On cherche à caractériser les entiers m possédant la propriété suivante :
(P) tout entier se terminant par m a son carré se terminant par m.

a. Soit $\ell \geqslant 1$ le nombre de chiffres de m. Montrer que m vérifie (P) si, et seulement si, $m^2 \equiv m\ [10^\ell]$.

b. Montrer qu'un entier $m \in [\![0, 10^\ell - 1]\!]$ possède la propriété (P) si, et seulement si, il existe un idempotent $e \in \mathbb{Z}_{10}$ tel que

$$\pi_{10,\ell-1}(e) = [m]_{\ell-1} \in \mathbb{Z}/10^\ell\mathbb{Z}.$$

Indication. Utiliser le lemme de Hensel.

c. Montrer que $\mathbb{Z}_{10}$ possède exactement quatre idempotents, et calculer tous les entiers m d'au plus six chiffres possédant la propriété (P).

Indication. Utiliser l'isomorphisme d'anneaux $\mathbb{Z}_{10} \simeq \mathbb{Z}_2 \times \mathbb{Z}_5$ de l'exercice 4.

E7. Démontrer les résultats énoncés dans la remarque 3.13.

E8. Soient p, ℓ deux nombres premiers distincts.

a. Montrer que le polynôme $X^2 - p \in \mathbb{Q}[X]$ est irréductible dans $\mathbb{Q}_p$. En déduire que les corps $\mathbb{Q}_p$ et $\mathbb{R}$ ne sont pas isomorphes.

b. Soient $u, v \in \mathbb{Z}$ tels que $up + v\ell = 1$. Montrer que $X^2 + upX + p\ell \in \mathbb{Q}[X]$ est irréductible dans $\mathbb{Q}_p$, mais possède une racine dans $\mathbb{Q}_\ell$. En déduire que les corps $\mathbb{Q}_p$ et $\mathbb{Q}_\ell$ ne sont pas isomorphes.

Indication. Utiliser le lemme de Hensel.

E9. Calculer le développement de Hensel de $-\frac{1}{9}$ dans $\mathbb{Q}_5$.

E10. Soit $\pi \in A$ un élément premier régulier. On suppose que $A/(\pi)$ est un corps. Soit $P \in A_\pi[X]$. On suppose qu'il existe $\alpha_0 \in A_\pi$, et $m \geqslant 0$ tel que $P(\alpha_0) \equiv 0\ [\pi^{m+1}]$ et $k = v_\pi\big(P'(\alpha_0)\big) \leqslant \dfrac{m}{2}$.

Montrer qu'il existe $\alpha \in \mathbb{Z}_p$ tel que $P(\alpha) = 0$ et $\alpha \equiv \alpha_0\ [\pi^{m-k+1}]$.

Indication. Montrer qu'il existe une suite $(x_n)_{n\geqslant 0}$ d'éléments de A_π telle que

(1) $x_0 = \alpha_0$, et $x_{n+1} \equiv x_n\ [\pi^{m-k+n+1}]$, pour tout $n \geqslant 0$;

(2) $P(x_n) \equiv 0\ [\pi^{m+n+1}]$.

E11. Soit p un nombre premier, et soit $x \in \mathbb{Q}_p^\times$. On écrit

$$x = up^m, \quad m \in \mathbb{Z},\ u \in \mathbb{Z}_p^\times.$$

a. On suppose que $p > 2$. Montrer que x est un carré dans $\mathbb{Q}_p$ si, et seulement si, m est pair et u est un carré modulo p.

b. Montrer que x est un carré dans $\mathbb{Q}_2$ si, et seulement si, l'entier m est pair et $u \equiv 1\ [8]$.

c. Montrer que l'on a des isomorphismes de groupes

$$\mathbb{Q}_2^\times/\mathbb{Q}_2^{\times 2} \simeq (\mathbb{Z}/2\mathbb{Z})^3 \quad \text{et} \quad \mathbb{Q}_p^\times/\mathbb{Q}_p^{\times 2} \simeq (\mathbb{Z}/2\mathbb{Z})^2, \quad \text{pour tout } p > 2.$$

E12. Soit $n \geqslant 2$ et soit p un nombre premier. On note $\mu_n(\mathbb{Q}_p)$ le groupe des racines n-ièmes de l'unité de $\mathbb{Q}_p$. On dira qu'une racine n-ième est primitive si elle engendre un sous-groupe de $\mathbb{Q}_p^\times$ d'ordre n.

a. Vérifier que $\mu_n(\mathbb{Q}_p) \subset \mathbb{Z}_p^\times$.

b. On suppose que n est premier à p. Montrer que la réduction modulo p induit un morphisme injectif

$$\mu_n(\mathbb{Q}_p) \hookrightarrow \mathbb{F}_p^\times.$$

En déduire que tout racine n-ième de l'unité est une racine $(p-1)$-ième de l'unité.

c. Montrer que $\mathbb{Q}_p^\times$ contient exactement $p-1$ racines $(p-1)$-ièmes de l'unité.

d. On suppose que $p > 2$. Pour tout $k \geqslant 1$, montrer que 1 est la seule racine p^k-ième de l'unité contenue dans $\mathbb{Q}_p$.

Indication. Montrer par récurrence que si $t^p = 1$, alors $t \equiv 1\ [p^{n+1}]$ pour tout $n \geqslant 0$.

e. Montrer que le polynôme $X^2 + 1 \in \mathbb{Q}_2[X]$ n'a pas de racines dans $\mathbb{Q}_2$, et en déduire que pour tout $k \geqslant 2$, $\mathbb{Q}_2$ ne contient pas de racines primitives 2^k-ièmes si $k \geqslant 2$.

f. En déduire que les seules racines de l'unité contenues dans $\mathbb{Q}_2$ sont ± 1, et que si $p > 2$, les seules racines de l'unité contenues dans $\mathbb{Q}_p$ sont les racines $(p-1)$-ièmes de l'unité.

E13. Soit p un nombre premier.

a. Pour tout $n \geqslant 0$, montrer que $v_p(n!) = \sum_{k \geqslant 1} E\left(\frac{n}{p^k}\right)$, où $E(x)$ désigne la partie entière d'un réel x.

b. Soit $a \in \mathbb{Z}_p$. On suppose que $v_p(a) \geqslant 1$ si $p > 2$ et $v_p(a) \geqslant 3$ si $p = 2$. Montrer que la série

$$9801 + \sum_{n \geqslant 1} \frac{\frac{1}{2} \times \left(\frac{1}{2} - 1\right) \times \cdots \times \left(\frac{1}{2} - n + 1\right)}{n!} a^n$$

converge vers l'unique racine carrée $\alpha \in \mathbb{Z}_p$ de $1 + a$ telle que $\alpha \equiv 1$ [p] si $p > 2$, et $\alpha \equiv 1$ [8] si $p = 2$.

c. Calculer le développement de Hensel de α lorsque $p = 7$ et $a = -\frac{7}{9}$.

Quatrième partie

Des corps ! Des corps partout !

Chapitre XXV

Constructions à la règle et au compas et théorie des corps

Si un algébriste est malade, est-ce faute d'anticorps ?

Dans ce chapitre, nous allons motiver l'étude approfondie de la théorie des corps en expliquant comment celle-ci apparaît naturellement dans les problèmes de construction à la règle et au compas, et en montrant comment elle permet de résoudre des problèmes datant de l'Antiquité, tels que la trisection de l'angle, la duplication du cube ou la quadrature du cercle.

1. Préliminaires

Commençons par mettre en place le contexte géométrique qui nous servira tout au long de ce chapitre.

Le problème qui va nous intéresser est le suivant : étant donné un segment-unité OI et éventuellement un nombre fini de points du plan, peut-on construire une figure donnée à partir de ces points à l'aide de la règle et du compas seulement ?

1.1. Définition. Soit $\mathscr{A}_0$ un ensemble fini de points du plan contenant les points O et I.

On dit qu'un point P du plan est *élémentairement constructible à la règle et au compas à partir de* $\mathscr{A}_0$, ou *élémentairement* $\mathscr{A}_0$*-constructible* si P

est obtenu par intersection de deux objets géométriques appartenant aux ensembles suivants :

(1) l'ensemble des droites passant par deux points distincts de $\mathscr{A}_0$;
(2) l'ensemble des cercles centrés en un point de $\mathscr{A}_0$ et dont le rayon est la distance entre deux points de $\mathscr{A}_0$.

On dit qu'un point P du plan est *constructible à la règle et au compas à partir de* $\mathscr{A}_0$ s'il existe un nombre fini de points $P_1, \ldots, P_r = P$ tels que, pour tout $i \in [\![1, r]\!]$, le point P_i est élémentairement constructible à partir de $\mathscr{A}_0 \cup \{P_1, \ldots, P_{i-1}\}$.

Une longueur ℓ est dite *constructible à la règle et au compas à partir de* $\mathscr{A}_0$ s'il existe deux points P, Q constructibles à la règle et au compas à partir de $\mathscr{A}_0$ tels que $PQ = \ell$.

Remarquons que, puisqu'un compas permet de reporter les longueurs, cela équivaut à demander l'existence d'un point P constructible à la règle et au compas à partir de $\mathscr{A}_0$ tel que $OP = \ell$.

Enfin, une droite du plan est dite *constructible à la règle et au compas à partir de* $\mathscr{A}_0$ si elle passe par deux points constructibles à la règle et au compas à partir de $\mathscr{A}_0$.

Dans ce qui suit, on abrégera souvent l'expression « constructible à la règle et au compas à partir de $\mathscr{A}_0$ » par l'expression « $\mathscr{A}_0$-constructible ». De plus, si $\mathscr{A}_0 = \{O, I\}$, on dira simplement « constructible », plutôt que « $\mathscr{A}_0$-constructible ».

Parmi les divers problèmes de construction à la règle et au compas les plus célèbres, on peut citer :

(i) la quadrature du cercle : étant donné un segment-unité, est-il possible de construire à la règle et au compas un carré dont l'aire est la même que celle du cercle-unité ? Autrement dit, la longueur $\sqrt{\pi}$ est-elle constructible ?
(ii) la duplication du cube : étant donné un segment-unité, peut-on construire l'arête d'un cube dont le volume est le double du cube-unité, à la règle et au compas ? Autrement dit, peut-on construire la longueur $\sqrt[3]{2}$?
(iii) la trisection de l'angle : étant donné un angle θ, peut-on construire l'angle $\dfrac{\theta}{3}$ à la règle et au compas ?
(iv) la construction du n-gone régulier : étant donné un entier $n \geqslant 2$, peut-on construire les sommets d'un n-gone régulier à la règle et au compas ?

Nous reviendrons sur ces considérations plus loin. Pour l'instant, commençons par des résultats plus modestes.

1.2. Lemme. *Soit $\mathscr{A}_0$ un ensemble fini de points du plan contenant O et I. Soient A, B et C des points $\mathscr{A}_0$-constructibles. On suppose que A et B sont distincts. Alors :*

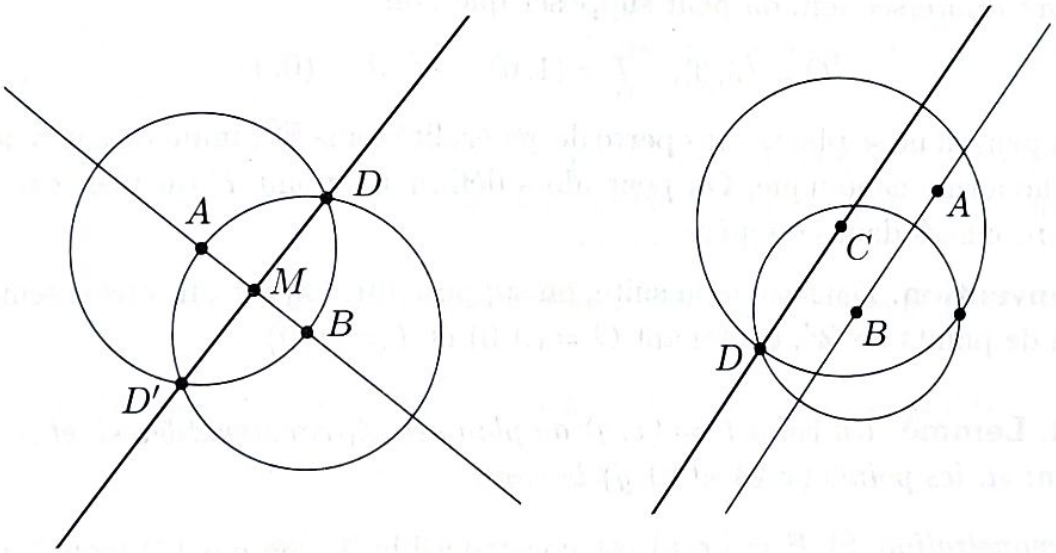
(1) *le milieu du segment AB et sa médiatrice sont $\mathscr{A}_0$-constructibles ;*

(2) *la parallèle à (AB) passant C est $\mathscr{A}_0$-constructible ;*

(3) *la perpendiculaire à (AB) passant C est $\mathscr{A}_0$-constructible.*

Démonstration

(1) Soient D et D' les points d'intersection des cercles de rayon AB centrés en A et B respectivement. Alors, (DD') est la médiatrice de AB, et son intersection M avec (AB) est le milieu de AB (figure de gauche).

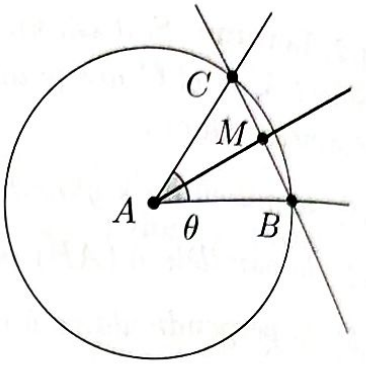


(2) Il n'y a rien à faire si $C = A$ ou $C = B$. On peut donc supposer que les points A, B, C sont deux à deux distincts. Le cercle de centre B et de rayon AC et le cercle de centre C et de rayon AB s'intersectent en deux points. Un des deux points d'intersection, que l'on note D, est tel que la droite (CD) soit parallèle à (AB) (figure de droite).

(3) En considérant l'intersection de (AB) et du cercle de centre A et de rayon AB, on obtient un point $B' \in (AB)$ tel que A soit le milieu de BB'. La perpendiculaire Δ à (AB) passant par A est alors la médiatrice de BB', qui est constructible par (1). La perpendiculaire voulue est alors l'unique parallèle à Δ passant par C, qui est constructible par (2) (puisque par définition Δ est obtenue à l'aide de deux points $\mathscr{A}_0$-constructibles). □

1.3. Corollaire. *Soit θ un angle formé par deux demi-droites du plan. Alors, la bissectrice de θ est constructible. Autrement dit, la bissection de l'angle est possible à la règle et au compas.*

Démonstration. Soit A le point d'intersection des deux demi-droites. En traçant le cercle de centre A et de rayon OI, on obtient deux points d'intersection B et C. Or, la médiatrice de BC est constructible d'après le lemme 1.2 (1) et elle passe par A. La bissectrice est alors la portion convenable de cette médiatrice. □



D'après le lemme 1.2, à partir de O et I, on peut construire la perpendiculaire à (OI) passant par O. En choisissant un point d'intersection J du cercle de centre O et de rayon OI avec cette perpendiculaire, on peut construire un repère orthonormé. Quitte à changer la numérotation et le point d'intersection, on peut supposer que l'on a

$$O = (0,0), \quad I = (1,0) \quad \text{et} \quad J = (0,1).$$

On peut donc se placer sans perte de généralité dans $\mathbb{R}^2$, muni de son repère orthonormé canonique. On peut alors définir un point P du plan par ses coordonnées dans ce repère.

Convention. Dans toute la suite, on suppose que $\mathscr{A}_0$ est un sous-ensemble fini de points de $\mathbb{R}^2$, contenant $O = (0,0)$ et $I = (1,0)$.

1.4. Lemme. *Un point $P = (x,y)$ du plan est $\mathscr{A}_0$-constructible si, et seulement si, les points $(x,0)$ et $(0,y)$ le sont.*

Démonstration. Si $P = (x,y)$ est constructible, le lemme 1.2 montre que la perpendiculaire et la parallèle à l'axe des abscisses passant par P sont constructibles. Leurs intersections avec l'axe des abscisses et l'axe des ordonnées sont respectivement $(x,0)$ et $(0,y)$. Inversement, si $Q = (x,0)$ et $R = (0,y)$ sont constructibles, ce même lemme montre que l'on peut construire la parallèle à l'axe des ordonnées passant par Q et la parallèle à l'axe des abscisses passant par R. L'intersection de ces deux droites est le point P. □

1.5. Remarque. Une utilisation élémentaire du compas montre que $(0,y)$ est $\mathscr{A}_0$-constructible si, et seulement si, $(y,0)$ l'est. Autrement dit, par le lemme précédent, $P = (x,y)$ est $\mathscr{A}_0$-constructible si, et seulement si, $(x,0)$ et $(y,0)$ le sont. Cela motive la définition suivante.

1.6. Définition. Un nombre réel $x \in \mathbb{R}$ est *constructible à la règle et au compas à partir de $\mathscr{A}_0$* si le point $(x,0)$ est constructible à la règle et au compas à partir de $\mathscr{A}_0$.

Un nombre complexe $z = x + iy \in \mathbb{C}$ est *constructible à la règle et au compas à partir de* $\mathscr{A}_0$ si le point (x, y) est constructible à la règle et au compas à partir de $\mathscr{A}_0$.

D'après la remarque précédente, cela équivaut au fait que les réels x, y sont constructibles à la règle et au compas à partir de $\mathscr{A}_0$.

Rappelons (cf. définition XV-3.11) que si L est un corps, un sous-corps de L est un sous-anneau K de L tel que pour tout $x \in K^\times$, on ait $x^{-1} \in K^\times$.

On peut alors établir un premier lien entre le problème de construction à la règle et au compas et la théorie des corps.

1.7. Théorème. *L'ensemble des nombres complexes $\mathscr{A}_0$-constructibles est un sous-corps de $\mathbb{C}$ contenant $\mathbb{Q}$.*

De plus, si $z \in \mathbb{C}$ est $\mathscr{A}_0$-constructible, toute racine carrée de z (i.e. tout nombre complexe de carré z) est aussi $\mathscr{A}_0$-constructible.

Démonstration. On va tout d'abord montrer que l'ensemble des nombres réels $\mathscr{A}_0$-constructibles est un sous-corps de $\mathbb{R}$ contenant $\mathbb{Q}$, et que de plus, si $r > 0$ est $\mathscr{A}_0$-constructible, alors $\sqrt{r}$ l'est également.

Bien entendu, 0 et 1 sont $\mathscr{A}_0$-constructibles, puisque $\mathscr{A}_0$ contient $O = (0, 0)$ et $I = (1, 0)$.

Soient maintenant $\alpha, \beta \in \mathbb{R}$ deux réels $\mathscr{A}_0$-constructibles. Les points

$$A = (\alpha, 0) \quad \text{et} \quad B = (\beta, 0)$$

sont donc $\mathscr{A}_0$-constructibles.

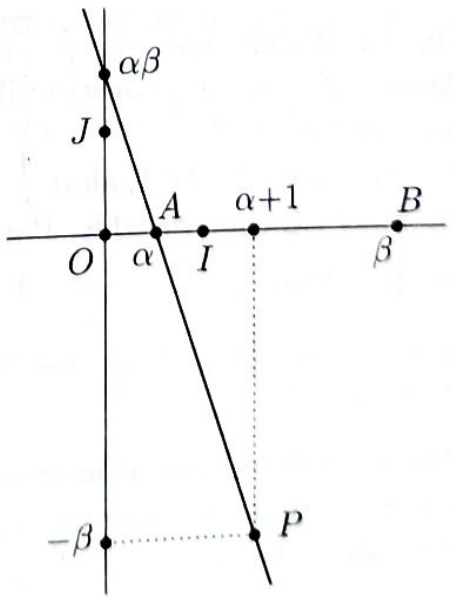
En reportant les longueurs grâce au compas, on construit l'unique point D de l'axe des abscisses tel que $\overrightarrow{BD} = \overrightarrow{AO}$. Autrement dit, on a $\overrightarrow{OD} = \overrightarrow{AB}$, et donc $D = (\beta - \alpha, 0)$. Ainsi, $\beta - \alpha$ est $\mathscr{A}_0$-constructible.

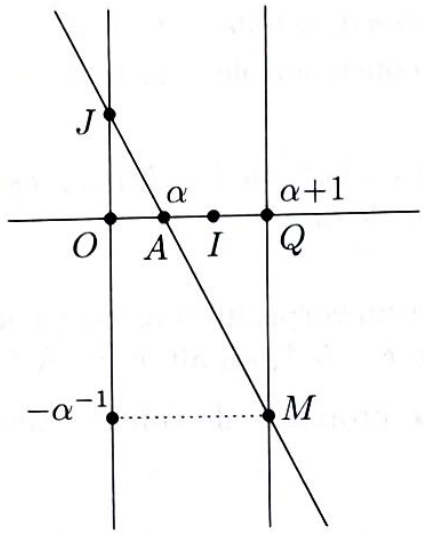
Passons au cas du produit. Comme $1, \alpha$ et β sont $\mathscr{A}_0$-constructibles, $\alpha + 1$ et $-\beta$ le sont aussi d'après ce qui précède, et le point $P = (\alpha + 1, -\beta)$ est donc $\mathscr{A}_0$-constructible.

On vérifie alors aisément que l'équation de la droite (AP) est donnée par

$$y = -\beta(x - \alpha).$$

Son intersection avec l'axe des ordonnées est le point $(0, \alpha\beta)$, on en déduit que le réel $\alpha\beta$ est $\mathscr{A}_0$-constructible conformément à la remarque 1.5.





Enfin, supposons que $\alpha \neq 0$. Comme le réel $\alpha + 1$ est $\mathscr{A}_0$-constructible, le point $Q = (\alpha+1, 0)$ est $\mathscr{A}_0$-constructible. De plus, le point $J = (0, 1)$ l'est aussi. L'équation de la droite (JA) est alors

$$y = -\alpha^{-1}(x - \alpha).$$

Son intersection avec la parallèle à l'axe des ordonnées passant par Q est donc le point $M = (\alpha+1, -\alpha^{-1})$. En particulier, le réel $-\alpha^{-1}$ est $\mathscr{A}_0$-constructible, et par conséquent α^{-1} l'est également.

Ainsi, l'ensemble des réels $\mathscr{A}_0$-constructibles est un sous-corps de $\mathbb{R}$.

Soient maintenant a, $b \in \mathbb{Z}$, avec $b \neq 0$. Puisque 1 est $\mathscr{A}_0$-constructible, -1 aussi, et une récurrence immédiate montre que a, $b \in \mathbb{Z}$ le sont également. Mais alors, $\frac{1}{b}$ l'est aussi. On en déduit alors que $\frac{a}{b}$ est $\mathscr{A}_0$-constructible.

Enfin, supposons que $r > 0$ soit $\mathscr{A}_0$-constructible.

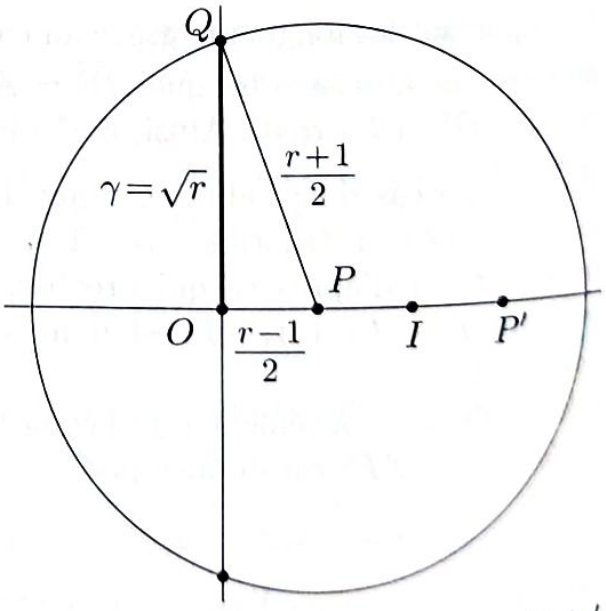
Supposons tout d'abord que $r > 1$. Puisque $1, 2$ et r sont $\mathscr{A}_0$-constructibles, les réels $\frac{r-1}{2}$ et $\frac{r+1}{2}$ le sont aussi.

Soient $P = \left(\frac{r-1}{2}, 0\right)$ et $P' = \left(\frac{r+1}{2}, 0\right)$. Soit $Q = (0, \gamma)$ le point d'intersection du cercle de centre P et de rayon OP' avec l'axe des ordonnées, tel que $\gamma > 0$. Un tel point existe, car $r > 1$. Le triangle OPQ étant rectangle en O, on a

$$\gamma^2 + \left(\frac{r-1}{2}\right)^2 = \left(\frac{r+1}{2}\right)^2.$$

On en déduit que $\gamma = \sqrt{r}$. Ainsi, $\sqrt{r}$ est $\mathscr{A}_0$-constructible dans ce cas.

Si maintenant $0<r<1$, alors $\frac{1}{r}>1$ est aussi $\mathscr{A}_0$-constructible. Par le cas précédent $\sqrt{\frac{1}{r}} = \frac{1}{\sqrt{r}}$ l'est aussi, et son inverse $\sqrt{r}$ est $\mathscr{A}_0$-constructible.



Passons maintenant au cas général. Supposons que $z = x+iy$ et $z' = x'+iy'$ soient $\mathscr{A}_0$-constructibles. Les réels x, x', y, y' sont donc $\mathscr{A}_0$-constructibles. Mais alors, les réels $x - x'$, $y - y'$, $xx' - yy'$ et $xy' + x'y$ le sont aussi.

Par conséquent, les nombres complexes

$$z - z' = (x - x') + i(y - y') \quad \text{et} \quad zz' = (xx' - yy') + i(xy' + x'y)$$

sont $\mathscr{A}_0$-constructibles.

De plus, si $z \neq 0$, $\dfrac{1}{x^2 + y^2}$ est $\mathscr{A}_0$-constructible, et par conséquent

$$z^{-1} = \frac{x - iy}{x^2 + y^2}$$

est encore $\mathscr{A}_0$-constructible.

Ainsi, l'ensemble des nombres complexes $\mathscr{A}_0$-constructibles est bien un sous-corps de $\mathbb{C}$, contenant $\mathbb{Q}$, d'après ce qui précède.

Montrons maintenant que si $z \in \mathbb{C}$ est $\mathscr{A}_0$-constructible, toute racine carrée de z l'est aussi. On peut toujours supposer $z \neq 0$. Écrivons alors $z = re^{i\theta}$, avec $r > 0$ et $\theta \in [0, 2\pi[$. Les racines carrées de z sont alors $\pm\sqrt{r}e^{i\frac{\theta}{2}}$. Soit P le point du plan d'affixe z. Alors, l'intersection du cercle de centre O et de rayon OP et de l'axe des abscisses est le point $(r, 0)$. Ainsi, le réel $r > 0$ est $\mathscr{A}_0$-constructible. D'après le cas précédent, $\sqrt{r}$ est $\mathscr{A}_0$-constructible. Autrement dit, $A = (\sqrt{r}, 0)$ est $\mathscr{A}_0$-constructible. En intersectant le cercle $\mathscr{C}$ de centre 0 et de rayon OA, et la droite (OP), on construit le point P' d'affixe $z' = \sqrt{r}e^{i\theta}$.

Les intersections de la médiatrice de AP' et du cercle $\mathscr{C}$ sont les points d'affixes $\pm\sqrt{r}e^{i\frac{\theta}{2}}$. Cela achève la démonstration. □

Nous allons maintenant introduire le langage des extensions de corps afin de poursuivre notre investigation.

2. Extensions de corps

2.1. Définition. Soit K un corps. Une *extension* de K est un corps L contenant K comme sous-corps. On le note L/K.

On vérifie aisément que L, muni de son produit et de la loi externe

$$K \times L \longrightarrow L$$
$$(\lambda, x) \longmapsto \lambda x,$$

est une K-algèbre associative unitaire.

On appelle *degré* de L/K la dimension de L comme K-espace vectoriel. On le note $[L : K]$.

Un *morphisme* d'extensions de K est un morphisme de K-algèbres.

Une *sous-extension* de L/K est une extension L'/K telle que L' soit un sous-corps de L.

Le lemme suivant sera abondamment utilisé dans la suite sans référence ultérieure.

2.2. Lemme. *Soient L, L' deux corps. Alors, tout morphisme $\varphi : L \longrightarrow L'$ est injectif. En particulier, tout morphisme d'extensions d'un même corps K est injectif.*

Démonstration. Gardons les notations de l'énoncé. Le noyau de $\varphi : L \longrightarrow L'$ est un idéal de L, donc est égal à (0) ou L. Dans le second cas, on aurait

$$1_{L'} = \varphi(1_L) = 0_{L'},$$

ce qui est impossible puisqu'un corps est non trivial. Ainsi, $\mathrm{Ker}(\varphi) = (0)$ et φ est injective. □

2.3. Remarque. Soit K un corps, et soit

$$\begin{aligned} \Theta_K : \mathbb{Z} &\longrightarrow K \\ n &\longmapsto n{\cdot}1_K \end{aligned}$$

le morphisme d'anneaux canonique.

Si $\mathrm{car}(K) = 0$, ce morphisme est injectif, et identifie le sous-anneau engendré par 1_K à $\mathbb{Z}$. Dans ce cas, puisque tout élément non nul de K est inversible, la proposition XV-3.18 montre que l'application

$$\begin{aligned} \mathbb{Q} &\longrightarrow K \\ \frac{a}{b} &\longmapsto (a{\cdot}1_K)(b{\cdot}1_K)^{-1} \end{aligned}$$

est un morphisme d'anneaux injectif bien défini. Ainsi, K contient un sous-corps isomorphe à $\mathbb{Q}$. Si $\mathrm{car}(K) = p$ (p nombre premier), le morphisme Θ_K se factorise en un morphisme d'anneaux injectif

$$\begin{aligned} \mathbb{F}_p &\longrightarrow K \\ \overline{n} &\longmapsto n{\cdot}1_K. \end{aligned}$$

Ainsi, K contient un sous-corps isomorphe à $\mathbb{F}_p$.

Autrement dit, tout corps K peut se voir comme une extension de $\mathbb{Q}$ ou de $\mathbb{F}_p$, à une identification canonique près.

On continue en étudiant les propriétés du degré.

2.4. Proposition (Formule de multiplicativité des degrés)

Soient M/K et L/M des extensions de corps. Alors, L/K est de degré fini si, et seulement si, M/K et L/M sont de degrés finis. Dans ce cas, on a

l'égalité

$$[L : K] = [L : M]\times[M : K].$$

En particulier, si L/K est de degré fini, toute sous-extension M/K est de degré fini, et l'on a

$$[M : K] \mid [L : K].$$

Démonstration. Soit $(e_i)_{i\in I}$ une K-base de M et soit $(f_j)_{j\in J}$ une M-base de L. Montrons que $(e_i f_j)_{(i,j)\in I\times J}$ est une K-base de L. Soit $x \in L$. La famille $(f_j)_{j\in J}$ étant M-génératrice, il existe une famille $(\lambda_j)_{j\in J}$ d'éléments de M à support fini telle que

$$x = \sum_{j\in J} \lambda_j f_j.$$

De même, pour tout $j \in J$, il existe $(\mu_{ij})_{i\in I}$ d'éléments de K à support fini telle que

$$\lambda_j = \sum_{i\in I} \mu_{ij} e_i.$$

Alors, la famille $(\mu_{ij})_{(i,j)\in I\times J}$ est à support fini et l'on a

$$x = \sum_{j\in J} \Big(\sum_{i\in I} \mu_{ij} e_i \Big) f_j = \sum_{(i,j)\in I\times J} \mu_{ij}(e_i f_j).$$

Ainsi, la famille $(e_i f_j)_{(i,j)\in I\times J}$ engendre L comme K-espace vectoriel. Montrons qu'elle est K-libre. Soit $(\mu_{ij})_{(i,j)\in I\times J}$ une famille d'éléments de K à support fini telle que

$$\sum_{(i,j)\in I\times J} \mu_{ij}(e_i f_j) = 0.$$

La famille $\big(\sum_{i\in I} \mu_{ij} e_i\big)_{j\in J}$ d'éléments de M est alors à support fini, et l'on a

$$\sum_{j\in J} \Big(\sum_{i\in I} \mu_{ij} e_i \Big) f_j = 0.$$

La famille $(f_j)_{j\in J}$ étant M-libre, on obtient

$$\sum_{i\in I} \mu_{ij} e_i = 0, \quad \text{pour tout } j \in J,$$

puis $\mu_{ij} = 0$ pour tout $(i, j) \in I \times J$, car $(e_i)_{i\in I}$ est K-libre.

Ainsi, la famille $(e_i f_j)_{(i,j)\in I\times J}$ est une K-base de L, équipotente à $I \times J$. Elle est alors finie si, et seulement si, I et J sont finis, c'est-à-dire si, et seulement si, $[L : M]$ et $[M : K]$ sont finis. Dans ce cas, on a

$$[L : K] = |I \times J| = |I|\times|J| = [L : M]\times[M : K].$$

Le dernier point est immédiat. □

On continue ce paragraphe en introduisant la notion de sous-extension engendrée par une partie. On vérifie facilement que l'intersection de sous-corps d'un même corps L est encore un sous-corps de L. Cela justifie la définition suivante.

2.5. Définition. Soit L/K une extension, et soit $\mathscr{S}$ une partie de L. On note $K(\mathscr{S})$ l'intersection des sous-corps de L contenant K et $\mathscr{S}$. C'est donc la plus petite sous-extension de L/K contenant $\mathscr{S}$.

L'extension $K(\mathscr{S})/K$ s'appelle la *sous-extension de L/K engendrée par $\mathscr{S}$*. Si $\mathscr{S} = \{s_1, \ldots, s_n\}$, on la note $K(s_1, \ldots, s_n)/K$.

On dit que L/K est *monogène* s'il existe $\alpha \in L$ tel que $L = K(\alpha)$.

2.6. Remarque. Il existe des extensions L/K de degré fini non monogènes (cf. exercice XXVI-3).

2.7. Proposition. *Soit L/K une extension de corps, et soient $\mathscr{S}$, $\mathscr{S}_1$, $\mathscr{S}_2$ des parties de L. Alors :*

(1) *pour $i = 1, 2$, on a $K(\mathscr{S}_i) \subset K(\mathscr{S}_1 \cup \mathscr{S}_2)$. De plus, on a les égalités entre extensions de $K(\mathscr{S}_1)$*

$$K(\mathscr{S}_1 \cup \mathscr{S}_2) = K(\mathscr{S}_1)(\mathscr{S}_2)$$

et les égalités entre extensions de $K(\mathscr{S}_2)$

$$K(\mathscr{S}_1 \cup \mathscr{S}_2) = K(\mathscr{S}_2)(\mathscr{S}_1).$$

En particulier, pour tous éléments $\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m \in L$, on a

$$\begin{aligned} K(\alpha_1, \ldots, \alpha_n, \beta_1, \ldots, \beta_m) &= K(\alpha_1, \ldots, \alpha_n)(\beta_1, \ldots, \beta_m) \\ &= K(\beta_1, \ldots, \beta_m)(\alpha_1, \ldots, \alpha_n); \end{aligned}$$

(2) *on a*

$$\begin{aligned} K(\mathscr{S}) = \big\{ & P(s_1, \ldots, s_n) Q(s_1, \ldots, s_n)^{-1} \mid \\ & n \geqslant 0,\ s_1, \ldots, s_n \in \mathscr{S},\ P, Q \in K[X_1, \ldots, X_n] \\ & \text{tels que } Q(s_1, \ldots, s_n) \neq 0 \big\}. \end{aligned}$$

En particulier, si $\mathscr{S} = \{s_1, \ldots, s_k\}$, alors on a

$$\begin{aligned} K(s_1, \ldots, s_k) = \big\{ & P(s_1, \ldots, s_k) Q(s_1, \ldots, s_k)^{-1} \mid P, Q \in K[X_1, \ldots, X_k] \\ & \text{tels que } Q(s_1, \ldots, s_k) \neq 0 \big\}. \end{aligned}$$

Démonstration

(1) Par définition, $K(\mathscr{S}_1 \cup \mathscr{S}_2)$ est un sous-corps de L contenant à la fois K et $\mathscr{S}_1 \cup \mathscr{S}_2$. Il contient donc K et S_i, d'où l'inclusion $K(\mathscr{S}_i) \subset K(\mathscr{S}_1 \cup \mathscr{S}_2)$ par définition de $K(\mathscr{S}_i)$.

De même, puisque $K(\mathscr{S}_1 \cup \mathscr{S}_2)$ contient $K(\mathscr{S}_1)$ et $\mathscr{S}_2$, on a aussi l'inclusion

$$K(\mathscr{S}_1)(\mathscr{S}_2) \subset K(\mathscr{S}_1 \cup \mathscr{S}_2).$$

Par ailleurs, comme $K(\mathscr{S}_1)(\mathscr{S}_2)$ contient $K(\mathscr{S}_1)$ et $\mathscr{S}_2$, il contient $K, \mathscr{S}_1$ et $\mathscr{S}_2$. C'est donc un sous-corps de L contenant K et $\mathscr{S}_1 \cup \mathscr{S}_2$. On a donc

$$K(\mathscr{S}_1 \cup \mathscr{S}_2) \subset K(\mathscr{S}_1)(\mathscr{S}_2),$$

d'où la première égalité. La seconde égalité se démontre en échangeant les rôles de $\mathscr{S}_1$ et $\mathscr{S}_2$, et le dernier point est clair.

(2) Notons L' le membre de droite dans l'égalité de l'énoncé. Soit $n \geqslant 0$, et soient $s_1, \ldots, s_n \in \mathscr{S} \subset L$. Comme L est en particulier un anneau contenant K, la stabilité par addition et par produit montre que le corps L contient $P(s_1, \ldots, s_n)$ pour tout $P \in K[X_1, \ldots, X_n]$. La stabilité par passage à l'inverse montre alors que $L' \subset K(\mathscr{S})$. Pour obtenir l'autre inclusion, il suffit de montrer que L' est un sous-corps de L contenant K et $\mathscr{S}$.

Le fait que L' contienne K et $\mathscr{S}$ est clair. En particulier, on a $0, 1 \in L'$. Soient x, y deux éléments de L'. Il existe donc deux entiers n, $m \geqslant 0$, ainsi que des polynômes $P_1, Q_1 \in K[X_1, \ldots, X_n]$ et $P_2, Q_2 \in K[X_{n+1}, \ldots, X_{n+m}]$, et des éléments $s_1, \ldots, s_{n+m} \in \mathscr{S}$ tels que

$$x = P_1(s_1, \ldots, s_n) Q_1(s_1, \ldots, s_n)^{-1}$$

et

$$y = P_2(s_{n+1}, \ldots, s_{n+m}) Q_2(s_{n+1}, \ldots, s_{n+m})^{-1}.$$

On définit trois polynômes P, Q, $R \in K[X_1, \ldots, X_{n+m}]$ par

$$P = P_1 Q_2 - P_2 Q_1, \quad Q = Q_1 Q_2, \quad R = P_1 P_2.$$

Par intégrité de L, on a

$$Q(s_1, \ldots, s_{n+m}) = Q_1(s_1, \ldots, s_n) Q_2(s_{n+1}, \ldots, s_{n+m}) \neq 0,$$

et l'on a alors sans difficulté les égalités

$$x - y = P(s_1, \ldots, s_{n+m}) Q(s_1, \ldots, s_{n+m})^{-1}$$

et

$$xy = R(s_1, \ldots, s_{n+m}) Q(s_1, \ldots, s_{n+m})^{-1}.$$

De plus, si $x \neq 0$, on a nécessairement $P_1(s_1, \ldots, s_n) \neq 0$, et alors

$$x^{-1} = Q_1(s_1, \ldots, s_n) P_1(s_1, \ldots, s_n)^{-1}.$$

Par conséquent, L' est bien un sous-corps de L, d'où l'égalité désirée.

Le reste provient du fait que si $\mathscr{S} = \{s_1, \ldots, s_k\}$, toute expression polynomiale en un nombre fini d'éléments de $\mathscr{S}$ peut se récrire comme une expression polynomiale en $s_1, \ldots, s_k$. Cela achève la démonstration. □

2.8. Exemples

(1) Soit L/K une extension. Si $\mathscr{S} \subset K$, on a $K(\mathscr{S}) = K$, puisque K est un sous-corps de L contenant $\mathscr{S}$.

(2) Soit L/K une extension. Pour tout $\alpha \in L$, et tous $a, b \in K$, $b \neq 0$, on a

$$K(a + b\alpha) = K(\alpha).$$

En effet, puisque $a + b\alpha \in K(\alpha)$, on a $K(a + b\alpha) \subset K(\alpha)$. Puisque b est non nul, on a aussi

$$\alpha = -b^{-1}a + b^{-1}(a + b\alpha) \in K(a + b\alpha),$$

d'où l'inclusion $K(\alpha) \subset K(a + b\alpha)$.

(3) On a l'égalité

$$\mathbb{Q}(i, \sqrt{2}) = \mathbb{Q}(\zeta_8),$$

où $\zeta_8 = e^{\frac{2i\pi}{8}} = \dfrac{\sqrt{2}}{2} + i\dfrac{\sqrt{2}}{2}$.

En effet, il est clair que $\zeta_8 \in \mathbb{Q}(i, \sqrt{2})$, d'où l'inclusion $\mathbb{Q}(\zeta_8) \subset \mathbb{Q}(i, \sqrt{2})$. D'autre part, on a $\zeta_8^2 = e^{i\frac{\pi}{2}} = i$, d'où $i \in \mathbb{Q}(\zeta_8)$. Mais alors, on a

$$\sqrt{2} = \frac{2\zeta_8}{1 + i} \in \mathbb{Q}(\zeta_8).$$

Il s'avère ainsi que i et $\sqrt{2} \in \mathbb{Q}(\zeta_8)$, d'où l'inclusion manquante.

(4) Soit K un sous-corps de $\mathbb{C}$, et soit $d \in K$. Deux racines carrées de d dans $\mathbb{C}$ différant seulement par un signe, elles engendrent la même sous-extension. Aussi, par commodité, on note $\sqrt{d} \in \mathbb{C}$ un nombre complexe de carré d.

Alors, on a l'égalité

$$K(\sqrt{d}) = \{a + b\sqrt{d} \mid a, b \in K\}.$$

En particulier, $[K(\sqrt{d}) : K] \leqslant 2$.

Si $\sqrt{d} \in K$, alors $K(\sqrt{d}) = K$, et il n'y a rien à faire. On peut donc supposer que $\sqrt{d} \notin K$. L'inclusion « $\supset$ » étant triviale, montrons l'autre inclusion. Il suffit pour cela de démontrer que l'ensemble

$$\{a + b\sqrt{d} \mid a, b \in K\}$$

est un sous-corps de $\mathbb{C}$ contenant K et $\sqrt{d}$. Tous les points à vérifier sont immédiats, sauf la stabilité par passage à l'inverse.

Supposons que $z = a + b\sqrt{d} \neq 0$. Alors, $a^2 - db^2 \neq 0$. En effet, si on avait $a^2 - db^2 = 0$, alors $b \neq 0$ (sinon $a = 0$ et $z = 0$), et par conséquent d serait le carré d'un élément de K, ce qui entraînerait

alors que $\sqrt{d} \in K$. Mais alors, on vérifie que l'on a

$$z^{-1} = \frac{a - b\sqrt{d}}{a^2 - db^2} \in \{a + b\sqrt{d} \mid a, b \in K\},$$

d'où le résultat.

Nous allons maintenant pouvoir résoudre (tout du moins en théorie) le problème de la construction à la règle et au compas. Commençons par introduire une nouvelle notation.

Notation. Si $\mathscr{A}$ est un ensemble fini de points du plan, on note $\mathbb{Q}(\mathscr{A})/\mathbb{Q}$ la sous-extension de $\mathbb{C}$ engendré par les coordonnées des éléments de $\mathscr{A}$.

2.9. Exemple. Si $\mathscr{A} = \{(0,0), (1,0)\}$, alors $\mathbb{Q}(\mathscr{A}) = \mathbb{Q}$.

On a alors le résultat suivant.

2.10. Lemme. *Le sous-corps des nombres complexes $\mathscr{A}_0$-constructibles est une extension de $\mathbb{Q}(\mathscr{A}_0)$.*

En particulier, tout élément de $\mathbb{Q}(\mathscr{A}_0)$ est $\mathscr{A}_0$-constructible.

Démonstration. Soit E l'ensemble des nombres complexes $\mathscr{A}_0$-constructibles. D'après le théorème 1.7, E est une extension de $\mathbb{Q}$. Par hypothèse, E contient les coordonnées des points de $\mathscr{A}_0$. Il contient donc également la sous-extension engendrée par celles-ci, à savoir $\mathbb{Q}(\mathscr{A}_0)$. □

Nous allons maintenant nous intéresser de plus près à la structure des extensions de $\mathbb{Q}(\mathscr{A}_0)$ obtenues à partir de points $\mathscr{A}_0$-constructibles. On commence par le cas des constructions élémentaires.

2.11. Proposition. *Soit P un point élémentairement $\mathscr{A}_0$-constructible. Alors, $\mathbb{Q}(\mathscr{A}_0 \cup \{P\})$ est une extension de $\mathbb{Q}(\mathscr{A}_0)$, et l'on a*

$$[\mathbb{Q}(\mathscr{A}_0 \cup \{P\}) : \mathbb{Q}(\mathscr{A}_0)] \leqslant 2.$$

Démonstration. Le premier point est clair, puisque $\mathscr{A}_0 \subset \mathscr{A}_0 \cup \{P\}$. Montrons le second point.

Soient $P_0 = (x_0, y_0)$, $P_1 = (x_1, y_1)$, $P_2 = (x_2, y_2)$ trois points de $\mathscr{A}_0$. Supposons que P_0 et P_1 soient distincts.

Alors :

(i) l'équation de la droite (P_0P_1) est donnée par

$$(x_1 - x_0)(y - y_0) - (y_1 - y_0)(x - x_0) = 0,$$

et est donc de la forme

$$ax + by + c = 0, \quad \text{avec } a,\ b,\ c \in \mathbb{Q}(\mathscr{A}_0)$$

(ii) l'équation du cercle de centre P_2 et de rayon P_0P_1 est donnée par

$$(x - x_2)^2 + (y - y_2)^2 = (x_1 - x_0)^2 + (y_1 - y_0)^2,$$

et est donc de la forme

$$x^2 + y^2 + ax + by + c = 0, \quad \text{avec } a,\ b,\ c \in \mathbb{Q}(\mathscr{A}_0).$$

Supposons tout d'abord que $P = (x, y)$ soit l'intersection de deux droites passant par des points de $\mathscr{A}_0$. D'après ce qui précède, (x, y) est l'unique solution d'un système linéaire de la forme

$$\begin{cases} ax + by + c = 0 \\ a'x + b'y + c' = 0, \end{cases}$$

avec $a,\ b,\ c,\ a', b', c' \in \mathbb{Q}(\mathscr{A}_0)$. En utilisant les formules de Cramer (ou en résolvant le système directement), on constate que x, y sont des fractions rationnelles en $a,\ a', b,\ b', c,\ c'$ à coefficients dans $\mathbb{Q}$, donc des éléments du corps $\mathbb{Q}(\mathscr{A}_0)$. Dans ce cas, on a

$$\mathbb{Q}(\mathscr{A}_0)(x, y) = \mathbb{Q}(\mathscr{A}_0),$$

puisque le plus petit sous-corps de $\mathbb{C}$ contenant $\mathbb{Q}(\mathscr{A}_0)$, x et y est $\mathbb{Q}(\mathscr{A}_0)$ lui-même.

En particulier, $[\mathbb{Q}(\mathscr{A}_0 \cup \{P\}) : \mathbb{Q}(\mathscr{A}_0)] = 1$.

Supposons maintenant que $P = (x, y)$ soit l'intersection d'une droite et d'un cercle du type précédent. On a ainsi

$$\begin{cases} ax + by + c = 0 \\ x^2 + y^2 + a'x + b'y + c' = 0, \end{cases}$$

avec $a,\ b,\ c,\ a', b', c' \in \mathbb{Q}(\mathscr{A}_0)$.

Remarquons que $(a, b) \neq (0, 0)$. Supposons par exemple que $b \neq 0$, et donc que $y = -\dfrac{ax + c}{b}$. Il s'ensuit que $y \in \mathbb{Q}(\mathscr{A}_0)(x)$. Par conséquent,

$$\mathbb{Q}(\mathscr{A}_0)(x, y) = \mathbb{Q}(\mathscr{A}_0)(x)(y) = \mathbb{Q}(\mathscr{A}_0)(x).$$

En remplaçant y par son expression dans la seconde équation, on obtient

$$ux^2 + vx + w = 0, \quad \text{avec } u,\ v,\ w \in \mathbb{Q}(\mathscr{A}_0).$$

On a nécessairement $(u, v) \neq (0, 0)$. Sinon, cela signifierait que l'intersection du cercle et de la droite est soit vide, ce qui n'est pas le cas par hypothèse,

soit infinie, ce qui est impossible. Supposons que $u = 0$. On a donc $v \neq 0$, et par conséquent $x = -wv^{-1} \in \mathbb{Q}(\mathscr{A}_0)$. On conclut alors comme précédemment. Si $u \neq 0$, on a

$$x = \frac{-v \pm \omega}{2u},$$

où ω^2 vérifie $\omega^2 = v^2 - 4uw$. D'après l'exemple 2.8 (2), on a

$$\mathbb{Q}(\mathscr{A}_0)(x) = \mathbb{Q}(\mathscr{A}_0)(\omega).$$

On utilise alors l'exemple 2.8 (4) pour conclure. Le cas $a \neq 0$ se traite de la même manière.

Finalement, supposons que $P = (x, y)$ soit l'intersection de deux cercles distincts du type précédent. Ainsi, (x, y) est solution d'un système de la forme

$$\begin{cases} x^2 + y^2 + ax + by + c = 0 \\ x^2 + y^2 + a'x + b'y + c' = 0, \end{cases}$$

avec a, b, c, $a', b', c' \in \mathbb{Q}(\mathscr{A}_0)$.

Remarquons que $(a, b) \neq (a', b')$, car sinon les deux cercles seraient égaux ou d'intersection vide. Ainsi, $(a - a', b - b') \neq (0, 0)$. En soustrayant la seconde équation à la première, on se ramène alors au cas précédent. Cela achève la démonstration. □

On peut alors facilement en déduire le résultat suivant.

2.12. Corollaire. *Soit $P = (x, y) \in \mathbb{R}^2$ un point $\mathscr{A}_0$-constructible. Alors, il existe une tour d'extensions*

$$E_0 = \mathbb{Q}(\mathscr{A}_0) \subset E_1 \subset \cdots \subset E_r$$

telle que

(1) *pour tout $i \in [\![1, r]\!]$, on a $[E_i : E_{i-1}] \leqslant 2$,*

(2) *on a x, $y \in E_r$.*

Démonstration. Soient $P_1, \ldots, P_r = P$ des points de $\mathbb{R}^2$ tels que, pour tout entier $i \in [\![1, r]\!]$, le point P_i soit élémentairement constructible à partir de l'ensemble $\mathscr{A}_0 \cup \{P_1, \ldots, P_{i-1}\}$.

Pour tout $i \in [\![0, r]\!]$, posons $E_i = \mathbb{Q}(\mathscr{A}_0 \cup \{P_1, \ldots, P_i\})$. Alors, on a bien entendu

$$E_0 = \mathbb{Q}(\mathscr{A}_0) \subset E_1 \subset \cdots \subset E_r,$$

et la proposition 2.11 montre que l'on a $[E_i : E_{i-1}] \leqslant 2$ pour tout $i \in [\![1, r]\!]$. Enfin, par définition, le corps E_r contient les coordonnées du point P, d'où le résultat. □

Ce corollaire conduit naturellement à poser la définition suivante.

2.13. Définition. On dit qu'une extension L/K est *2-décomposable* s'il existe une tour d'extensions

$$E_0 = K \subset E_1 \subset \cdots \subset E_r = L$$

vérifiant $[E_i : E_{i-1}] \in \{1, 2\}$, pour tout $i \in [\![1, r]\!]$.

2.14. Remarque. Le degré d'une extension 2-décomposable est fini et est une puissance de 2. En effet, la formule de multiplicativité des degrés montre que $[L : K]$ est le produit des degrés des extensions E_i/E_{i-1}.

On verra en revanche dans l'exercice 15 qu'une extension dont le degré est une puissance de 2 n'est pas nécessairement 2-décomposable.

2.15. Corollaire. *Tout nombre complexe $\mathscr{A}_0$-constructible est contenu dans une extension 2-décomposable de $\mathbb{Q}(\mathscr{A}_0)$.*

Démonstration. Soit $z = x + iy$ un nombre complexe $\mathscr{A}_0$-constructible. Alors, par définition, le point $P = (x, y)$ est $\mathscr{A}_0$-constructible. Par le corollaire 2.12, il existe une extension $L/\mathbb{Q}(\mathscr{A}_0)$ 2-décomposable contenant x et y. D'après l'exemple 2.8 (4), on a $[L(i) : L] \leqslant 2$. L'extension $L(i)/\mathbb{Q}(\mathscr{A}_0)$ est alors aussi 2-décomposable. De plus, $L(i)$ contient x, y et i, et contient donc z. □

2.16. Remarque. La démonstration du corollaire 2.12 montre en fait que l'extension 2-décomposable de $\mathbb{Q}(\mathscr{A}_0)$ contenant les coordonnées d'un point élémentairement constructible est un sous-corps de $\mathbb{R}$ (puisqu'elle est construite en rajoutant les coordonnées de points intermédiaires). On a donc en fait le résultat suivant.

Si $z \in \mathbb{C}$ est $\mathscr{A}_0$-constructible, il existe un sous-corps L de $\mathbb{R}$ tel que l'extension $L/\mathbb{Q}(\mathscr{A}_0)$ soit 2-décomposable, et tel que

(1) $z \in L$ si $z \in \mathbb{R}$;

(2) $z \in L(i)$ si $z \in \mathbb{C} \setminus \mathbb{R}$.

Nous en arrivons au théorème fondamental de la construction à la règle et au compas, qui donne une réciproque au corollaire 2.15.

2.17. Théorème. *Un nombre complexe est $\mathscr{A}_0$-constructible si, et seulement si, il est contenu dans une extension 2-décomposable de $\mathbb{Q}(\mathscr{A}_0)$.*

Démonstration. Le sens direct étant donné par le corollaire 2.15, il reste donc à voir que tout élément d'une extension 2-décomposable de $\mathbb{Q}(\mathscr{A}_0)$ est $\mathscr{A}_0$-constructible.

On commence par démontrer le résultat suivant : soit L/K une extension de degré 1 ou 2, où L et K sont des sous-corps de $\mathbb{C}$. Alors, si tout élément de K est $\mathscr{A}_0$-constructible, il en est de même de tout élément de L.

Soit $z \in L$. Puisque L est de dimension au plus 2 comme K-espace vectoriel, la famille $(1, z, z^2)$ est K-liée. Il existe donc u, v, $w \in K$ non tous nuls tels que

$$uz^2 + vz + w = 0.$$

De plus, on a $(u, v) \neq (0, 0)$, car sinon, on aurait $w = 0$.

Supposons tout d'abord que $u = 0$. On a donc $v \neq 0$, et ainsi $z = -wv^{-1}$ est $\mathscr{A}_0$-constructible par le théorème 1.7. Si $u \neq 0$, on a

$$z = \frac{-v \pm \omega}{2u},$$

où ω^2 vérifie $\omega^2 = v^2 - 4uw$. D'après ce même théorème, le nombre complexe ω est $\mathscr{A}_0$-constructible et par conséquent, z l'est aussi puisque u, v et ω le sont, d'où le résultat voulu.

Soit maintenant $L/\mathbb{Q}(\mathscr{A}_0)$ une extension 2-décomposable. Il existe donc une tour d'extensions

$$E_0 = \mathbb{Q}(\mathscr{A}_0) \subset E_1 \subset \cdots \subset E_r = L$$

vérifiant $[E_i : E_{i-1}] \leqslant 2$, pour tout entier $i \in [\![1, r]\!]$. Puisque tout élément de $\mathbb{Q}(\mathscr{A}_0)$ est $\mathscr{A}_0$-constructible par le lemme 2.10, le point précédent et une récurrence immédiate permettent de conclure. □

2.18. Corollaire (Critère de Wantzel). *Si $z \in \mathbb{C}$ est $\mathscr{A}_0$-constructible, alors l'extension $\mathbb{Q}(\mathscr{A}_0)(z)/\mathbb{Q}(\mathscr{A}_0)$ est de degré fini et $[\mathbb{Q}(\mathscr{A}_0)(z) : \mathbb{Q}(\mathscr{A}_0)]$ est une puissance de 2.*

Démonstration. D'après le théorème précédent, z est contenu dans un extension $L/\mathbb{Q}(\mathscr{A}_0)$ 2-décomposable. On a donc

$$\mathbb{Q}(\mathscr{A}_0) \subset \mathbb{Q}(\mathscr{A}_0)(z) \subset L.$$

La formule de multiplicativité des degrés montre alors que le degré de l'extension $\mathbb{Q}(\mathscr{A}_0)(z)/\mathbb{Q}(\mathscr{A}_0)$ divise le degré de $L/\mathbb{Q}(\mathscr{A}_0)$, qui est une puissance de 2, par la remarque 2.14. Le résultat en découle aisément. □

Attention ! La réciproque de ce résultat est fausse. On peut construire une infinité de nombres complexes non constructibles, mais engendrant une extension de degré 4 sur $\mathbb{Q}$ (cf. exercice 15).

Le critère de Wantzel amène une question cruciale : étant donnés une extension L/K et un élément $\alpha \in L$, comment savoir si $K(\alpha)/K$ est de degré fini, et dans l'affirmative, comment le calculer ?

Remarquons que si $K(\alpha)/K$ est de degré fini d, alors la famille $(1, \alpha, \ldots, \alpha^d)$ d'éléments de $K(\alpha)$ est liée, puisqu'elle possède $d+1$ éléments. Ainsi, il existe $a_0, \ldots, a_d \in K$ non tous nuls, tels que

$$a_d\alpha^d + \cdots + a_1\alpha + a_0 = 0.$$

Autrement dit, $\alpha \in L$ est racine d'un polynôme non nul à coefficients dans K. Cela conduit à la notion d'élément algébrique sur K.

3. Extensions algébriques, transcendantes

Si L/K est une extension de corps, et si $\alpha \in L$, il existe un unique morphisme de K-algèbres $ev_\alpha : K[X] \longrightarrow L$ tel que $ev_\alpha(X) = \alpha$. Si $P \in K[X]$, l'image de P est notée $P(\alpha)$. Si $P = a_nX^n + \cdots + a_1X + a_0$, on a donc

$$P(\alpha) = a_n\alpha^n + \cdots + a_1\alpha + a_0.$$

On note $K[\alpha]$ l'image de ev_α.

On a alors la définition suivante.

3.1. Définition. Soit L/K une extension de corps, et soit $\alpha \in L$. On dit que α est *algébrique sur* K s'il existe un polynôme $P \in K[X]$ non nul tel que $P(\alpha) = 0$. On dit que α est *transcendant sur* K sinon.

On dit que L/K est *algébrique* si tout élément de L est algébrique sur K. On dit que L/K est *transcendante* sinon.

3.2. Remarque. Soit L/K une extension, et soit M/K une sous-extension. Si L/K est algébrique, alors M/K et L/M sont algébriques.

En effet, si tout élément de L est algébrique sur K, c'est a fortiori vrai pour tout élément de M, et M/K est donc une extension algébrique. De plus, tout polynôme à coefficients dans K peut être vu comme un polynôme à coefficients dans M. Ainsi, tout élément algébrique sur K l'est aussi sur M. En particulier, si L/K est algébrique, L/M est algébrique.

3.3. Exemples

(1) Charles Hermite et Carl Louis Ferdinand von Lindemann ont respectivement démontré que e et π sont transcendants sur $\mathbb{Q}$.

(2) En revanche, le réel π est algébrique sur $\mathbb{R}$, puisqu'il est racine du polynôme $X - \pi \in \mathbb{R}[X]$.

(3) Le nombre complexe $i \in \mathbb{C}$ est algébrique sur $\mathbb{Q}$, puisque $i^2 + 1 = 0$.

(4) Si T est une indéterminée sur K, alors T est transcendant sur K.

(5) Si L/K est une extension de degré fini, alors tout élément $\alpha \in L$ est algébrique sur K.

En effet, si $d = [L : K] = \dim_K(L)$, alors la famille $(1, \alpha, \ldots, \alpha^d)$ est liée, puisqu'elle possède $d+1$ éléments. Ainsi, il existe $a_0, \ldots, a_d \in K$ non tous nuls, tels que $a_d\alpha^d + \cdots + a_1\alpha + a_0 = 0$. Par conséquent, le polynôme $P = a_dX^d + \cdots + a_1X + a_0 \in K[X]$ est non nul et vérifie $P(\alpha) = 0$.

3.4. Remarque. On montre facilement que l'ensemble des nombres complexes algébriques sur $\mathbb{Q}$ est dénombrable. En particulier, si l'on prend un nombre complexe « au hasard », il y a de fortes chances qu'il soit transcendant. Néanmoins, il est extrêmement difficile en pratique de déterminer si un nombre complexe donné est algébrique ou transcendant.

3.5. Proposition. *Soit L/K une extension, et soit $\alpha \in L$. Alors, l'ensemble*

$$I_\alpha = \{P \in K[X] \mid P(\alpha) = 0\}$$

est un idéal de $K[X]$. Il est non nul si, et seulement si, α est algébrique sur K. Dans ce cas, il existe un unique polynôme unitaire irréductible $\mu_{\alpha,K}$ de $K[X]$ tel que

$$\mu_{\alpha,K}(\alpha) = 0.$$

Ce polynôme est l'unique générateur unitaire de I_α.

Démonstration. Par définition, I_α est le noyau du morphisme d'évaluation

$$ev_\alpha : K[X] \longrightarrow L.$$

C'est donc un idéal de $K[X]$. Il est non nul si, et seulement si, α est algébrique, par définition même. Dans ce cas, il est engendré par un polynôme non nul, car $K[X]$ est principal. De plus, comme $K[X]$ est intègre, deux générateurs d'un même idéal diffèrent par un élément inversible. On en déduit facilement qu'il existe un unique générateur unitaire $\mu_{\alpha,K}$. De plus, il est non nul et non inversible, car non constant.

Le théorème de factorisation donne un isomorphisme entre $K[X]/(\mu_{\alpha,K})$ et un sous-anneau de l'anneau intègre L. Ainsi, $K[X]/(\mu_{\alpha,K})$ est intègre, et le polynôme $\mu_{\alpha,K}$ est irréductible par la proposition XXI-1.9.

Si $f \in K[X]$ est un polynôme irréductible, unitaire et vérifiant $f(\alpha) = 0$, alors $f \in I_\alpha = (\mu_{\alpha,K})$, et donc $\mu_{\alpha,K} \mid f$. Comme ces deux polynômes sont irréductibles unitaires, ils sont égaux. □

3.6. Définition. Le polynôme $\mu_{\alpha,K}$ est appelé le *polynôme minimal de α sur K*.

3.7. Exemples

(1) Soit $K = \mathbb{R}$, et soit $\alpha = i$. Alors, $f = X^2 + 1 \in \mathbb{R}[X]$ annule α. Il est unitaire, et irréductible, car il est de degré 2 et n'a pas de racines réelles. Ainsi, $\mu_{\alpha,\mathbb{R}} = X^2 + 1$.

(2) Soit $K = \mathbb{Q}$ et soit $\alpha = \sqrt[4]{3}$. Alors, $f = X^4 - 3 \in \mathbb{Q}[X]$ annule α. Il est unitaire, à coefficients dans $\mathbb{Z}$, donc primitif, et irréductible dans $\mathbb{Z}[X]$, et donc dans $\mathbb{Q}[X]$, par le critère d'Eisenstein. Ainsi, $\mu_{\alpha,\mathbb{Q}} = X^4 - 3$.

(3) Soit $K = \mathbb{Q}$, et soit $\alpha = \sqrt{2} + \sqrt{3}$. On vérifie que $f = X^4 - 10X^2 + 1$ annule α. Il est irréductible par la proposition XXI-4.4.
Ainsi, $\mu_{\alpha,\mathbb{Q}} = X^4 - 10X^2 + 1$.

3.8. Théorème. *Soit L/K une extension. Alors, $\alpha \in L$ est algébrique sur K si, et seulement si, $K(\alpha)/K$ est une extension de degré fini. Dans ce cas, une K-base de $K(\alpha)$ est $(1, \alpha, \ldots, \alpha^{d-1})$, où $d = \deg(\mu_{\alpha,K})$. En particulier, on a $K(\alpha) = K[\alpha]$, ainsi que*

$$[K(\alpha) : K] = \deg(\mu_{\alpha,K}).$$

Démonstration. Si $K(\alpha)/K$ est de degré fini, alors $\alpha \in K(\alpha)$ est algébrique sur K par l'exemple 3.3 (5). Si α est algébrique, le morphisme d'évaluation donne un isomorphisme de K-algèbres

$$K[X]/(\mu_{\alpha,K}) \simeq K[\alpha].$$

Comme un isomorphisme d'espaces vectoriels envoie une base sur une base, on déduit de la proposition XXI-1.9 que la famille $(1, \alpha, \ldots, \alpha^{d-1})$ est une K-base de $K[\alpha]$. Par cette même proposition, $K[X]/(\mu_{\alpha,K})$ est un corps. Ainsi, $K[\alpha]$ est aussi un corps, et il contient K comme sous-corps. Par conséquent, $K[\alpha]/K$ est une sous-extension de L/K contenant α, et l'on a donc $K(\alpha) \subset K[\alpha]$. Mais, on a aussi $K[\alpha] \subset K(\alpha)$, d'où $K[\alpha] = K(\alpha)$. L'extension $K(\alpha)/K$ est donc de degré d, de base $(1, \alpha, \ldots, \alpha^{d-1})$. □

3.9. Remarque. Si $\alpha \in L$ est algébrique sur K, la démonstration du théorème nous fournit un isomorphisme de K-algèbres

$$K[X]/(\mu_{\alpha,K}) \xrightarrow{\sim} K(\alpha),$$

qui envoie $\overline{X}$ sur α.

3.10. Exemples

(1) D'après les exemples 3.7, on a

$$[\mathbb{R}(i) : \mathbb{R}] = 2, \quad [\mathbb{Q}(\sqrt[4]{3}) : \mathbb{Q}] = 4, \quad \text{et} \quad [\mathbb{Q}(\sqrt{2} + \sqrt{3}) : \mathbb{Q}] = 4.$$

(2) Soit $d \in \mathbb{Z}$ non nul et sans facteurs carrés. Alors, $X^2 - d \in \mathbb{Q}[X]$ est irréductible puisqu'il est de degré 2 et n'a pas de racines. Comme il est unitaire et annule $\sqrt{d}$, on a $\mu_{\sqrt{d},\mathbb{Q}} = X^2 - d$. Le théorème précédent nous donne alors

$$\mathbb{Q}(\sqrt{d}) = \{a + b\sqrt{d} \mid a, b \in \mathbb{Q}\},$$

ainsi que $[\mathbb{Q}(\sqrt{d}) : \mathbb{Q}] = 2$.

3.11. Corollaire. *Soit L/K une extension. Si $\alpha_1, \ldots, \alpha_n \in L$ sont algébriques sur K, alors $K(\alpha_1, \ldots, \alpha_n)/K$ est de degré fini. En particulier, l'extension $K(\alpha_1, \ldots, \alpha_n)/K$ est algébrique, et l'on a*

$$K(\alpha_1, \ldots, \alpha_n) = K[\alpha_1, \ldots, \alpha_n],$$

où $K[\alpha_1, \ldots, \alpha_n] = \{P(\alpha_1, \ldots, \alpha_n) \mid P \in K[X_1, \ldots, X_n]\}$.

Démonstration. Pour tout $i \in [\![0, n]\!]$, on pose $E_i = K(\alpha_1, \ldots, \alpha_i)$. On a alors $E_i = E_{i-1}(\alpha_i)$ pour tout $i \in [\![1, n]\!]$. Puisque α_i est algébrique sur K, il est algébrique sur E_{i-1} (cf. remarque 3.2). Le théorème 3.8 montre alors que $[E_i : E_{i-1}]$ est fini, et la formule de multiplicativité des degrés implique que $[E_n : K]$ est fini, ce qu'il fallait vérifier.

D'après l'exemple 3.3 (5), $K(\alpha_1, \ldots, \alpha_n)/K$ est algébrique. Il reste à démontrer la dernière égalité. On procède par récurrence sur n. C'est trivialement vrai pour $n = 0$. Supposons que l'égalité soit vraie pour $n - 1$ éléments algébriques, et montrons-le pour n éléments. Gardons les notations précédentes. Comme $K(\alpha_1, \ldots, \alpha_n) = E_{n-1}(\alpha_n)$, et que α_n est algébrique sur E_{n-1}, on a $E_{n-1}(\alpha_n) = E_{n-1}[\alpha_n]$, d'après le théorème 3.8. Par hypothèse de récurrence, on a $E_{n-1} = K[\alpha_1, \ldots, \alpha_{n-1}]$, et ainsi

$$K(\alpha_1, \ldots, \alpha_n) = K[\alpha_1, \ldots, \alpha_{n-1}][\alpha_n] = K[\alpha_1, \ldots, \alpha_n],$$

la dernière égalité étant claire. Cela achève la démonstration. □

La définition suivante sera bien commode pour raccourcir les énoncés.

3.12. Définition. Soit L/K une extension, et soit $\alpha \in L$ un élément algébrique sur K. Le *degré* de α sur K est le degré de l'extension $K(\alpha)/K$. C'est donc aussi le degré de $\mu_{\alpha,K}$.

3.13. Corollaire. *Soit L/K une extension. Alors, l'ensemble des éléments de L qui sont algébriques sur K est une sous-extension de L/K.*

Démonstration. Soit M l'ensemble des éléments de L algébriques sur K. Il est clair que $K \subset M$. Soient $\alpha, \beta \in M$. Le corollaire 3.11 montre que $K(\alpha, \beta)/K$ est algébrique. En particulier, $-\alpha$, $\alpha + \beta$, $\alpha\beta \in K(\alpha, \beta)$

sont algébriques sur K. De même, si α est non nul, $\alpha^{-1} \in K(\alpha, \beta)$ est algébrique sur K. Cela achève la démonstration. □

3.14. Corollaire. *Soit L/K une extension, et soit $\mathscr{S}$ une partie de L. Alors, $K(\mathscr{S})/K$ est algébrique si, et seulement si, tout élément de $\mathscr{S}$ est algébrique sur K.*

Démonstration. Soit M l'ensemble des éléments de L qui sont algébriques sur K. Supposons que tout élément de $\mathscr{S}$ soit algébrique sur K. Alors, $\mathscr{S}$ est contenue dans M. Comme M est un sous-corps de L contenant K d'après le corollaire précédent, on a $K(\mathscr{S}) \subset M$. Autrement dit, $K(\mathscr{S})/K$ est algébrique. L'autre implication est immédiate. □

Avant de continuer, nous avons besoin d'une nouvelle définition.

3.15. Définition. Soit L/K une extension. Si E/K et E'/K sont deux sous-extensions de L/K, le *compositum* de E/K et E'/K dans L/K est la sous-extension de L/K engendrée par $E \cup E'$. Autrement dit, c'est la plus petite sous-extension de L/K contenant E et E'. On la note EE'/K.

3.16. Exemple. Si $E = K(\mathscr{S})$, on vérifie aisément que l'on a

$$EE' = E'(\mathscr{S}).$$

En particulier, on a

$$K(\mathscr{S})K(\mathscr{S}') = K(\mathscr{S} \cup \mathscr{S}').$$

3.17. Corollaire. *Soit L/K une extension, et soient E/K et E'/K deux sous-extensions de L/K. Si E/K est algébrique, alors EE'/E' est algébrique.*

Démonstration. Puisque tout élément de E est algébrique sur K, tout élément de E est aussi algébrique sur E' par la remarque 3.2. D'après le corollaire précédent, tout élément de la sous-extension M/E' de L/E' engendrée par E est algébrique. Mais cette extension n'est rien d'autre que EE'. En effet, M est un sous-corps de L contenant E et E'. Il contient aussi K, et par suite, il contient EE'. Inversement, EE' est une extension de E' contenant E par définition, et contient donc M, d'où finalement l'égalité $M = EE'$. Cela achève la démonstration. □

3.18. Corollaire. *Soit L/K une extension, et soient E/K et E'/K deux sous-extensions de L/K. Si E/K et E'/K sont algébriques, EE'/K est algébrique.*

Démonstration. Pour obtenir le résultat voulu, il suffit d'appliquer le corollaire 3.14 à $\mathscr{S} = E \cup E'$. □

On a également la proposition suivante.

3.19. Proposition. *Soit $K \subset M \subset L$ une tour d'extensions. Alors, L/K est algébrique si, et seulement si, M/K et L/M sont algébriques.*

Démonstration. Le sens direct est donné par la remarque 3.2. Supposons que M/K et L/M soient algébriques et soit $\alpha \in L$. Comme α est algébrique sur M par hypothèse, il existe $a_0, \ldots, a_n \in M$ non tous nuls tels que

$$a_n\alpha^n + \cdots + a_1\alpha + a_0 = 0.$$

Soit $K' = K(a_0, \ldots, a_n)$. L'égalité précédente montre en fait que α est algébrique sur K'. En particulier, $K'(\alpha)/K'$ est une extension de degré fini par le théorème 3.8. De plus, par hypothèse, $a_0, \ldots, a_n$ sont algébriques sur K. Par le corollaire 3.11, le degré $[K' : K]$ est fini. La formule de multiplicativité des degrés montre alors que $[K'(\alpha) : K]$ est fini. Comme on a $K \subset K(\alpha) \subset K'(\alpha)$, ce même résultat implique que $[K(\alpha) : K]$ est également fini. Le théorème 3.8 montre alors que α est algébrique sur K. Cela achève la démonstration. □

Le théorème 3.8 nous permet donc de calculer le degré d'une extension algébrique $K(\alpha)/K$, sous réserve que l'on sache calculer $\mu_{\alpha,K}$. En utilisant ce théorème de façon répétée, on peut aussi calculer le degré d'une extension de la forme $K(\alpha_1, \ldots, \alpha_n)/K$, où $\alpha_1, \ldots, \alpha_n$ sont algébriques sur K, mais les calculs peuvent devenir assez délicats, car le calcul d'un polynôme minimal sur un corps quelconque n'est pas toujours aisé.

Nous finissons ce paragraphe par quelques considérations sur ce sujet. Nous allons pour cela introduire la notion d'extensions linéairement disjointes.

Nous allons maintenant nous intéresser au problème suivant. On se fixe une fois pour toutes une extension L/K, et deux sous-extensions E/K et E'/K de L/K. Peut-on comparer les degrés de EE'/E' et de E/K ?

Un premier pas dans cette direction est donné par le lemme suivant.

3.20. Lemme. *Si E/K est algébrique, toute famille K-génératrice de E est une famille E'-génératrice de EE'. En particulier, on a*

$$EE' = \Big\{ \sum_{i=1}^{n} x_i x_i' \mid n \geqslant 0,\ x_i \in E,\ x_i' \in E' \Big\},$$

et l'on a

$$[EE' : E'] \leqslant [E : K].$$

Démonstration. Soit $\mathscr{S}$ une famille K-génératrice de E. En particulier, on a l'égalité $E = K(\mathscr{S})$, et ainsi $EE' = E'(\mathscr{S})$, par l'exemple 3.16.

Soit $y \in EE'$. D'après la proposition 2.7, y est une fraction rationnelle à coefficients dans E' en un nombre fini d'éléments $s_1, \ldots, s_n \in \mathscr{S}$. En particulier, $y \in E'(s_1, \ldots, s_n)$. Comme $s_1, \ldots, s_n$ sont algébriques sur K par hypothèse, donc algébriques aussi sur E', on a

$$E'(s_1, \ldots, s_n) = E'[s_1, \ldots, s_n],$$

d'après le corollaire 3.11. Or, un monôme en $s_1, \ldots, s_n$ est un élément de E. Il est donc combinaison K-linéaire d'éléments de $\mathscr{S}$. On en déduit que tout élément de $E'[s_1, \ldots, s_n]$ est une combinaison E'-linéaire d'éléments de $\mathscr{S}$. En particulier, c'est le cas de y. Ainsi, $\mathscr{S}$ est bien une famille E'-génératrice de EE'. L'égalité décrivant EE' est alors claire.

L'inégalité $[EE' : E'] \leqslant [E : K]$ s'obtient alors en appliquant ce qui précède à une K-base de E (même si E/K est de degré infini, car l'inégalité est alors triviale). Cela achève la démonstration. □

3.21. Remarque. Ce résultat est faux si l'extension E/K n'est pas algébrique. En effet, soient X, Y deux indéterminées sur $\mathbb{C}$, et soit $L = \mathbb{C}(X, Y)$. Posons $E = \mathbb{C}(X)$ et $E' = \mathbb{C}(Y)$. Alors, le théorème de décomposition en éléments simples montre que les éléments

$$X^n, \quad \frac{1}{(X-a)^m}, \quad n \geqslant 0,\ m \geqslant 1$$

engendrent E. En revanche, ces éléments n'engendrent pas $EE' = \mathbb{C}(X, Y)$ sur E', comme on peut le constater en considérant $\dfrac{1}{X+Y}$.

Nous allons maintenant nous intéresser à la conservation de la liberté d'une famille. On commence par un lemme.

3.22. Lemme. *Soient E/K et E'/K deux sous-extensions de L/K. Alors, les propriétés suivantes sont équivalentes :*

(1) *toute famille K-libre de E est encore une famille E'-libre de EE' ;*

(2) *toute famille K-libre de E' est encore une famille E-libre de EE'.*

Dans ce cas, on a nécessairement $E \cap E' = K$.

Démonstration. Il suffit de démontrer l'implication « (1) $\Longrightarrow$ (2) », l'autre implication s'en déduisant en échangeant les rôles de E et E'.

Il suffit bien entendu de considérer les familles libres de cardinal fini, vu la définition d'une famille libre. Soit $(x'_1, \ldots, x'_n)$ une famille K-libre d'éléments de E'. Supposons qu'elle soit E-liée. En particulier, $n \geqslant 1$. Il existe donc $x_1, \ldots, x_n \in E$ non tous nuls tels que

$$x_1 x'_1 + \cdots + x_n x'_n = 0.$$

Si $(x_1, \ldots, x_n)$ était K-libre, elle serait aussi E'-libre par hypothèse, et l'on aurait donc $x'_1 = \cdots = x'_n = 0$, ce qui contredirait la liberté de $(x'_1, \ldots, x'_n)$.

Ainsi, $(x_1, \ldots, x_n)$ est K-liée. Quitte à renuméroter, on peut supposer que la famille $(x_1, \ldots, x_r)$ est une base du sous-espace engendré par $(x_1, \ldots, x_n)$, pour un certain entier $r \in [\![1, n-1]\!]$.

Pour tout $j \in [\![r+1, n]\!]$, on peut alors écrire

$$x_j = \sum_{i=1}^{r} \lambda_{ij} x_i, \quad \lambda_{ij} \in K.$$

Mais alors, on a

$$x_1 x_1' + \cdots + x_r x_r' + \sum_{j=r+1}^{n} x_j' \Big(\sum_{i=1}^{r} \lambda_{ij} x_i \Big) = 0,$$

soit encore

$$\sum_{i=1}^{r} \Big(x_i' + \sum_{j=r+1}^{n} \lambda_{ij} x_j' \Big) x_i = 0.$$

Puisque $(x_1, \ldots, x_r)$ est K-libre, elle est E'-libre par hypothèse, d'où

$$x_i' = - \sum_{j=r+1}^{n} \lambda_{ij} x_j', \quad \text{pour tout } i \in [\![1, r]\!].$$

Mais alors, $(x_1', \ldots, x_n')$ est K-liée, d'où une contradiction. Cela démontre le résultat voulu.

Montrons la dernière partie. Supposons que la propriété (1) soit vérifiée, par exemple, mais que $E \cap E' \neq K$. En particulier, $[E \cap E' : K] \geqslant 2$. Il existe donc deux éléments x, $y \in E \cap E'$ tels que la famille (x, y) soit libre. En particulier, ces éléments sont tous deux non nuls. Or, deux tels éléments sont E'-liés, puisque $-yx + xy = 0$, d'où une contradiction. On a donc bien $E \cap E' = K$. □

3.23. Définition. Soit L/K une extension de corps.

Deux sous-extensions E/K et E'/K de L/K sont dites *linéairement disjointes* si elles vérifient une des conditions équivalentes du lemme précédent.

Dans ce cas, on a nécessairement $E \cap E' = K$.

3.24. Exemple. La condition $E \cap E' = K$ est nécessaire pour avoir la disjonction linéaire, mais elle n'est pas suffisante.

Posons en effet $E = \mathbb{Q}(\sqrt[3]{2})$ et $E' = \mathbb{Q}(j\sqrt[3]{2})$. Le polynôme unitaire $X^3 - 2$ est irréductible dans $\mathbb{Q}[X]$ d'après le critère d'Eisenstein. On a donc

$$\mu_{\sqrt[3]{2}, \mathbb{Q}} = X^3 - 2,$$

et ainsi $[E : \mathbb{Q}] = 3$. Puisque $E \cap E'/\mathbb{Q}$ est une sous-extension de $E/\mathbb{Q}$, on a $[E \cap E' : \mathbb{Q}] \mid [E : \mathbb{Q}]$. Ainsi, $[E \cap E' : \mathbb{Q}] = 1$ ou 3.

Si l'on avait $[E \cap E' : \mathbb{Q}] = 3$, on aurait $E \cap E' = E$ et donc $E' = E \subset \mathbb{R}$, ce qui n'est pas le cas. Par conséquent, $[E \cap E' : \mathbb{Q}] = 1$, ou encore $E \cap E' = \mathbb{Q}$.

Par ailleurs, $(1, \sqrt[3]{2}, (\sqrt[3]{2})^2)$ est une $\mathbb{Q}$-base de E. Elle est donc $\mathbb{Q}$-libre. En revanche, elle est E'-liée, puisque

$$2{\cdot}1 + (j\sqrt[3]{2})^2{\cdot}\sqrt[3]{2} + j\sqrt[3]{2}{\cdot}(\sqrt[3]{2})^2 = 2(1 + j + j^2) = 0.$$

L'intérêt de cette notion est donnée par la proposition suivante.

3.25. Proposition. *Soient E/K et E'/K deux sous-extensions de L/K. On suppose que E/K est de degré fini. Alors, on a*

$$[EE' : E'] \leqslant [E : K],$$

et l'on a égalité si, et seulement si, E/K et E'/K sont linéairement disjointes.

Démonstration. Puisqu'une extension de corps de degré fini est algébrique, l'inégalité $[EE' : E'] \leqslant [E : K]$ découle du lemme 3.20. Si les extensions E/K et E'/K sont linéairement disjointes, toute K-base de E reste une E'-base de EE', d'après la définition et ce même lemme. On a donc égalité dans ce cas. Inversement, supposons que l'on ait

$$[EE' : E'] = [E : K] = d.$$

Soit $(x_1, \ldots, x_r)$ une famille K-libre de E (donc $r \leqslant d$), que l'on complète en une K-base $(x_1, \ldots, x_d)$ de E. C'est donc une famille E'-génératrice de EE' par le lemme 3.20. Puisque EE' est de dimension d sur E' par hypothèse, c'est une E'-base de EE'. En particulier, la famille $(x_1, \ldots, x_r)$ est E'-libre. Par conséquent, E/K et E'/K sont linéairement disjointes. □

3.26. Corollaire. *Soient E/K et E'/K deux sous-extensions de L/K de degré fini. Alors, on a*

$$[EE' : E'] \leqslant [E : K],\ \ [EE' : E] \leqslant [E' : K],\ \ [EE' : K] \leqslant [E : K] \times [E' : K],$$

et l'on a $[EE' : K] = [E : K] \times [E' : K]$ si, et seulement si, E/K et E'/K sont linéairement disjointes.

En particulier, c'est le cas si $[E : K]$ et $[E' : K]$ sont premiers entre eux.

Démonstration. Les deux premières inégalités découlent de la proposition précédente. On a alors

$$[EE' : K] = [EE' : E'] \times [E' : K] \leqslant [E : K] \times [E' : K].$$

On a égalité si, et seulement si, $[EE' : E'] = [E : K]$, c'est-à-dire si, et seulement si, E/K et E'/K sont linéairement disjointes, toujours d'après cette même proposition.

Supposons que les entiers $[E : K]$ et $[E' : K]$ soient premiers entre eux. Puisque $[E : K]$ et $[E' : K]$ divisent $[EE' : K]$, le produit $[E : K] \times [E' : K]$ divise également $[EE' : K]$. En particulier, on a $[EE' : K] \geqslant [E : K] \times [E' : K]$.

Le premier point implique alors que $[EE' : K] = [E : K] \times [E' : K]$. Cela achève la démonstration. □

3.27. Exemple. Puisque $\alpha = \sqrt[4]{2}$ et $\beta = \sqrt[5]{2}$ sont algébriques sur $\mathbb{Q}$, l'extension $\mathbb{Q}(\alpha, \beta)/\mathbb{Q}$ est algébrique. D'après le critère d'Eisenstein, les polynômes $X^4 - 2$ et $X^5 - 2$ sont irréductibles dans $\mathbb{Q}[X]$. On a donc

$$[\mathbb{Q}(\alpha) : \mathbb{Q}] = 4 \quad \text{et} \quad [\mathbb{Q}(\beta) : \mathbb{Q}] = 5.$$

D'après le corollaire précédent, les extensions $\mathbb{Q}(\alpha)/\mathbb{Q}$ et $\mathbb{Q}(\beta)/\mathbb{Q}$ sont linéairement disjointes. On a alors

$$[\mathbb{Q}(\alpha, \beta) : \mathbb{Q}] = [\mathbb{Q}(\alpha) : \mathbb{Q}] \times [\mathbb{Q}(\beta) : \mathbb{Q}] = 20.$$

Nous allons maintenant pouvoir résoudre les problèmes de construction cités en début de chapitre.

4. Quelques problèmes de construction à la règle et au compas

Commençons par reformuler le critère de Wantzel dans le langage des extensions algébriques.

4.1. Corollaire (Critère de Wantzel, version 2.0). *Un nombre complexe $\mathscr{A}_0$-constructible est algébrique sur $\mathbb{Q}(\mathscr{A}_0)$, et son degré est une puissance de 2.*

Rappelons aussi que lorsque $\mathscr{A}_0 = \{O, I\}$, on a $\mathbb{Q}(\mathscr{A}_0) = \mathbb{Q}$.

Intéressons-nous maintenant à quelques problèmes de construction à la règle et au compas. On commence par le plus célèbre d'entre eux.

La quadrature du cercle. Comme expliqué au début de ce chapitre, il s'agit de savoir si $\sqrt{\pi}$ est constructible à la règle au compas à partir du segment-unité. Si c'était le cas, $\pi = (\sqrt{\pi})^2$ serait aussi constructible. En particulier, le critère de Wantzel entraînerait que $\mathbb{Q}(\pi)/\mathbb{Q}$ serait de degré fini, et donc que π serait algébrique sur $\mathbb{Q}$. Or, un célèbre théorème de Lindemann affirme que π est transcendant sur $\mathbb{Q}$. La quadrature du cercle est donc impossible.

La duplication du cube-unité. Cela revient à savoir si $\sqrt[3]{2}$ est constructible à partir du segment-unité. Comme on l'a déjà constaté, $\sqrt[3]{2}$ est algébrique sur $\mathbb{Q}$, de degré 3. D'après le critère de Wantzel, $\sqrt[3]{2}$ n'est pas constructible, et la duplication du cube-unité s'avère impossible.

La trisection de l'angle. On se donne un segment-unité OI, ainsi que trois points A, B, C du plan, tels que les deux demi-droites issues de A et passant respectivement par B et C forment un angle de mesure $\theta \in [0, \pi[$, non orienté.

En utilisant pour origine d'un repère orthonormé le point A, et une des deux droites (AB) ou (AC) comme un des axes, et en traçant le cercle de centre A et de rayon $OI = 1$, on peut toujours se ramener à la question suivante : peut-on construire le point $Q = \big(\cos\big(\frac{\theta}{3}\big), \sin\big(\frac{\theta}{3}\big)\big)$ à partir des points O, I et du point $P = \big(\cos(\theta), \sin(\theta)\big)$?

Remarquons qu'un point $\big(\cos(\varphi), \sin(\varphi)\big)$, avec $\varphi \in [0, \pi[$, est constructible à partir d'un ensemble $\mathscr{A}_0$ si, et seulement si, $\big(\cos(\varphi), 0\big)$ l'est, puisque l'on peut retrouver $\big(\cos(\varphi), \sin(\varphi)\big)$ en intersectant la droite $x = \cos(\varphi)$ avec la moitié supérieure du cercle-unité. Le problème de la trisection de l'angle se reformule donc de la façon suivante : soit $\mathscr{A}_0 = \{(0,0), (1,0), \big(\cos(\theta), 0\big)\}$. Le réel $\cos\big(\frac{\theta}{3}\big)$ est-il $\mathscr{A}_0$-constructible ?

Remarquons que dans ce cas, on a $\mathbb{Q}(\mathscr{A}_0) = \mathbb{Q}\big(\cos(\theta)\big)$. Pour tout $\varphi \in \mathbb{R}$, on a

$$e^{3i\varphi} = (e^{i\varphi})^3 = \big(\cos(\varphi) + i\sin(\varphi)\big)^3.$$

En développant et en identifiant les parties réelles, on obtient

$$\cos(3\varphi) = \cos(\varphi)^3 - 3\cos(\varphi)\sin(\varphi)^2 = 4\cos(\varphi)^3 - 3\cos(\varphi).$$

En particulier, le polynôme $P = 4X^3 - 3X - \cos(\theta) \in \mathbb{Q}\big(\cos(\theta)\big)[X]$ annule le réel $\alpha = \cos\big(\frac{\theta}{3}\big)$. Ainsi, $\mu_{\alpha,\mathbb{Q}(\cos(\theta))} \mid P$.

Si P possède une racine dans $\mathbb{Q}\big(\cos(\theta)\big)$, il n'est pas irréductible et par conséquent $\mu_{\alpha,\mathbb{Q}(\cos(\theta))}$ est de degré 1 ou 2. Ainsi, $\mathbb{Q}\big(\cos(\theta)\big)(\alpha)/\mathbb{Q}\big(\cos(\theta)\big)$ est 2-décomposable et α est $\mathscr{A}_0$-constructible par le théorème 2.17.

Si P n'a pas de racine dans $\mathbb{Q}\big(\cos(\theta)\big)$, alors P est irréductible puisque P est de degré 3. On a alors $\mu_{\alpha,\mathbb{Q}(\cos(\theta))} = \frac{1}{4}P$, et α est alors alors de degré 3 sur $\mathbb{Q}\big(\cos(\theta)\big)$. Dans ce cas, α n'est pas $\mathscr{A}_0$-constructible.

En résumé, nous avons démontré le résultat suivant.

4.2. Théorème. *Soit $\theta \in [0, \pi[$. Alors, l'angle de mesure θ est trisectable à la règle et au compas si, et seulement si, le polynôme*

$$4X^3 - 3X - \cos(\theta)$$

admet une racine dans $\mathbb{Q}\big(\cos(\theta)\big)$.

4.3. Exemples

(1) Soit $\theta = \frac{2\pi}{3}$. Alors, $\cos(\theta) = -\frac{1}{2}$, et ainsi $\mathbb{Q}\big(\cos(\theta)\big) = \mathbb{Q}$. On doit vérifier si le polynôme $4X^3 - 3X + \frac{1}{2}$ a une racine rationnelle ou non. Cela revient à savoir si $8X^3 - 6X + 1$ a une racine rationnelle ou non. Les racines rationnelles possibles sont ici

$$\pm 1, \quad \pm\frac{1}{2}, \quad \pm\frac{1}{4}, \quad \pm\frac{1}{8},$$

et l'on vérifie qu'aucune de ces possibilités ne convient. Ainsi, $\frac{2\pi}{3}$ n'est pas trisectable à la règle et au compas.

(2) Soit $\theta = \frac{\pi}{4}$. Ici, $\cos(\theta) = \frac{\sqrt{2}}{2}$, d'où $\mathbb{Q}\big(\cos(\theta)\big) = \mathbb{Q}(\sqrt{2})$. Il faut donc voir si le polynôme $4X^3 - 3X + \frac{\sqrt{2}}{2}$ possède une racine dans $\mathbb{Q}(\sqrt{2})$. On vérifie effectivement que $-\frac{\sqrt{2}}{2}$ est une telle racine. Par conséquent, l'angle de mesure $\frac{\pi}{4}$ est trisectable à la règle et au compas.
Il y a en fait une manière plus simple de le constater. Comme on a l'égalité $\cos\big(\frac{\pi}{6}\big) = \frac{\sqrt{3}}{2}$, on peut construire l'angle de mesure $\frac{\pi}{6}$ à la règle et au compas à partir du segment-unité, puis obtenir celui de mesure $\frac{\pi}{12}$ par bissection.

La construction du n-gone régulier. Le problème est ici de construire les sommets d'un n-gone régulier à la règle et au compas. On se convainc facilement qu'il suffit de savoir construire un n-gone régulier centré en O, et dont un des sommets est le point $I = (1, 0)$. Dans ce cas, il faut et il suffit de savoir construire le point $P = \big(\cos\big(\frac{2\pi}{n}\big), \sin\big(\frac{2\pi}{n}\big)\big)$. Une fois ce point construit, il suffit d'utiliser le compas pour obtenir les autres sommets en reportant la longueur IP.

Par conséquent, la question devient : étant donné un entier $n \geqslant 3$, le nombre complexe $\zeta_n = \cos\big(\frac{2\pi}{n}\big) + i \sin\big(\frac{2\pi}{n}\big)$ est-il constructible ?

Il faut donc examiner les propriétés de l'extension $\mathbb{Q}(\zeta_n)/\mathbb{Q}$.

Par définition, ζ_n est racine du n-ième polynôme cyclotomique Φ_n (cf. définition XXI-5.1). Ce polynôme est unitaire, irréductible dans $\mathbb{Q}[X]$ d'après le théorème XXI-5.9, et de degré $\varphi(n)$.

On a donc $\mu_{\zeta_n,\mathbb{Q}} = \Phi_n$ et $[\mathbb{Q}(\zeta_n) : \mathbb{Q}] = \varphi(n)$. Pour que ζ_n soit constructible, il faut donc que $\varphi(n)$ soit une puissance de 2. Or, on a

$$\varphi(n) = \prod_{p|n} p^{v_p(n)-1}(p-1).$$

On en déduit donc que, pour tout diviseur premier p impair de n, $p-1$ est une puissance de 2, et l'on a $v_p(n) = 1$.

Remarquons également que si $p = 1 + 2^k$ est un nombre premier, alors k une puissance de 2. En effet, supposons que k soit divisible par un nombre premier impair ℓ, et écrivons $k = \ell m$, avec $m \geqslant 1$. Alors, on a

$$p = 1 + (2^m)^\ell = (1 + 2^m)(2^{m(\ell-1)} - 2^{m(\ell-2)} + \cdots + 2^{2m} - 2^m + 1).$$

Ainsi, $1 + 2^m \mid p$. Puisque $\ell > 1$, $1 + 2^m < p$ et est est donc un diviseur strict de p, ce qui contredit la primalité de p.

4.4. Définition. Soit $m \geqslant 0$. L'entier $F_m = 1 + 2^{2^m}$ est appelé le *m-ième nombre de Fermat.*

On vient donc de démontrer le résultat partiel suivant.

4.5. Proposition. *Si le n-gone régulier est constructible, alors n est produit d'une puissance de 2 et de nombres de Fermat premiers deux à deux distincts.*

4.6. Remarque. Fermat conjecturait que les entiers F_m étaient tous premiers. C'est vrai pour $m = 0$, 1, 2, 3 et 4, puisque

$$F_0 = 3, \quad F_1 = 5, \quad F_2 = 17, \quad F_3 = 257 \quad \text{et} \quad F_4 = 65537,$$

mais faux pour $m = 5$, puisque

$$F_5 = 641 \cdot 6700417.$$

Les seuls nombres de Fermat premiers connus à ce jour sont $F_0, \ldots, F_4$. On ne sait pas à l'heure actuelle s'il existe une infinité de nombres de Fermat premiers.

La question qui se pose maintenant est de savoir si la réciproque de la proposition précédente est vraie. Commençons par réduire le problème au cas d'un nombre de côtés premier impair.

4.7. Lemme. *Soient $m, n \geqslant 1$ deux entiers premiers entre eux.*

Alors, le mn-gone régulier est constructible si, et seulement si, le n-gone régulier et le m-gone régulier sont constructibles.

Démonstration. Il suffit de voir que ζ_{mn} est constructible si, et seulement si, ζ_m et ζ_n le sont. Si ζ_{mn} est constructible, alors $\zeta_n = \zeta_{mn}^m$ et $\zeta_m = \zeta_{mn}^n$ sont aussi constructibles.

Inversement, supposons que ζ_m et ζ_n soient constructibles. Puisque m et n sont premiers entre eux, il existe $u, v \in \mathbb{Z}$ tels que $um + vn = 1$.

Mais alors, on a

$$\zeta_{mn} = (\zeta_{mn}^m)^u (\zeta_{mn}^n)^v = \zeta_n^u \zeta_m^v.$$

Ainsi, ζ_{mn} est aussi constructible. □

4.8. Remarque. Ce lemme peut aussi se démontrer par des considérations purement géométriques.

Remarquons que par bissections successives de l'angle plat, qui est évidemment constructible, on peut obtenir l'angle de mesure $\frac{2\pi}{2^m}$ pour tout entier $m \geqslant 0$. Ainsi, le polygone régulier à 2^m côtés est constructible pour tout entier $m \geqslant 0$. Il résulte de ce qui précède que la réciproque à la proposition 4.5 est vraie si, et seulement si, pour tout nombre premier de Fermat p, le p-gone régulier est constructible.

Nous allons maintenant démontrer que c'est effectivement le cas. **ZE** bonne idée[1] est de considérer le groupe des automorphismes de $\mathbb{Q}(\zeta_p)/\mathbb{Q}$. Avant d'expliquer pourquoi, introduisons une définition.

4.9. Définition. Soit L/K une extension. Le *groupe de Galois* de L/K est le groupe des automorphismes de l'extension L/K, c'est-à-dire le groupe des automorphismes de K-algèbres de L. On le note $\mathrm{Gal}(L/K)$.

Si $\sigma, \tau \in \mathrm{Gal}(L/K)$, on écrira $\sigma\tau$ au lieu de $\sigma \circ \tau$.

On a alors le lemme suivant.

4.10. Lemme. *Soit L un corps, et soit $\mathscr{S}$ un ensemble de morphismes d'anneaux $L \longrightarrow L$. Alors :*

(1) *l'ensemble*

$$L^{\mathscr{S}} = \{x \in L \mid \sigma(x) = x, \quad \text{pour tout } \sigma \in \mathscr{S}\}$$

est un sous-corps de L ;

(2) *si tous les éléments de $\mathscr{S}$ sont des automorphismes, on a*

$$L^{\mathscr{S}} = L^{\langle \mathscr{S} \rangle},$$

où $\langle \mathscr{S} \rangle$ est le sous-groupe de $\mathrm{Aut}(L)$ engendré par $\mathscr{S}$;

(3) *si L est une extension de K, et si $\mathscr{S} \subset \mathrm{Gal}(L/K)$, alors $L^{\mathscr{S}}$ est une extension de K.*

1. Due à Évariste Galois.

Démonstration. Montrons le premier point. Puisque les éléments de $\mathscr{S}$ sont des morphismes d'anneaux, on a $\sigma(0) = 0$ et $\sigma(1) = 1$, pour tout $\sigma \in \mathscr{S}$. Ainsi, $0, 1 \in L^{\mathscr{S}}$. Si de plus x, $y \in L^{\mathscr{S}}$, pour tout $\sigma \in \mathscr{S}$, on a

$$\sigma(x-y) = \sigma(x) - \sigma(y) = x - y \quad \text{et} \quad \sigma(xy) = \sigma(x)\sigma(y) = xy.$$

Par suite, $x - y$, $xy \in L^{\mathscr{S}}$. Enfin, si $x \in L^{\times}$, on a aussi

$$\sigma(x^{-1}) = \sigma(x)^{-1} = x^{-1}, \quad \text{pour tout } \sigma \in L^{\mathscr{S}},$$

et ainsi $x^{-1} \in L^{\mathscr{S}}$, d'où le premier point.

Supposons que $\mathscr{S} \subset \mathrm{Aut}(L)$. Comme $\mathscr{S} \subset \langle \mathscr{S} \rangle$, on a $L^{\mathscr{S}} \supset L^{\langle \mathscr{S} \rangle}$. Si maintenant $x \in L$ est fixé par tous les éléments de $\mathscr{S}$, il est fixé par les produits des éléments de $\mathscr{S}$ et de leurs inverses. On a donc l'inclusion manquante.

Le dernier point découle directement de (1) et du fait qu'un élément du groupe $\mathrm{Gal}(L/K)$ fixe K point par point. Cela achève la démonstration.□

Le groupe de Galois permet donc de construire des sous-extensions de L/K. Nous allons maintenant regarder le cas de l'extension $\mathbb{Q}(\zeta_n)/\mathbb{Q}$.

4.11. Lemme. *Soit $n \geqslant 2$ un entier. Pour tout entier $m \in \mathbb{Z}$ premier à n, il existe un unique élément de $\sigma_m \in \mathrm{Gal}\big(\mathbb{Q}(\zeta_n)/\mathbb{Q}\big)$ tel que $\sigma_m(\zeta_n) = \zeta_n^m$.*

De plus, l'application

$$\begin{array}{ccc} (\mathbb{Z}/n\mathbb{Z})^{\times} & \longrightarrow & \mathrm{Gal}\big(\mathbb{Q}(\zeta_n)/\mathbb{Q}\big) \\ \overline{m} & \longmapsto & \sigma_m \end{array}$$

est bien définie, et est un isomorphisme de groupes.

Autrement dit, $\mathrm{Gal}\big(\mathbb{Q}(\zeta_n)/\mathbb{Q}\big) \simeq (\mathbb{Z}/n\mathbb{Z})^{\times}$.

Démonstration. Soit $n \geqslant 2$ un entier, et soit $m \in \mathbb{Z}$ un entier premier à n.

Montrons l'existence de σ_m. Puisque les racines du n-ième polynôme cyclotomique Φ_n sont les racines primitives n-ièmes de l'unité, le morphisme d'évaluation

$$\begin{array}{ccc} \mathbb{Q}[X] & \longrightarrow & \mathbb{Q}(\zeta_n) \\ P & \longmapsto & P(\zeta_n^m) \end{array}$$

se factorise en un morphisme de $\mathbb{Q}$-algèbres

$$\begin{array}{ccc} \mathbb{Q}[X]/(\Phi_n) & \longrightarrow & \mathbb{Q}(\zeta_n) \\ \overline{P} & \longmapsto & P(\zeta_n^m). \end{array}$$

En utilisant la remarque 3.9, on obtient un morphisme de $\mathbb{Q}$-algèbres

$$\sigma_m : \mathbb{Q}(\zeta_n) \longrightarrow \mathbb{Q}(\zeta_n)$$

tel que $$\sigma_m\big(P(\zeta_n)\big) = P(\zeta_n^m), \quad \text{pour tout } P \in \mathbb{Q}[X].$$

C'est en particulier un endomorphisme injectif d'un $\mathbb{Q}$-espace vectoriel de dimension finie ; il est donc bijectif. Ainsi, $\sigma_m \in \mathrm{Gal}\big(\mathbb{Q}(\zeta_n)/\mathbb{Q}\big)$. Il vérifie par définition $\sigma_m(\zeta_n) = \zeta_n^m$.

Si maintenant $\sigma \in \mathrm{Gal}\big(\mathbb{Q}(\zeta_n)/\mathbb{Q}\big)$, puisque σ est $\mathbb{Q}$-linéaire et respecte le produit, on a

$$\sigma\big(P(\zeta_n)\big) = P\big(\sigma(\zeta_n)\big), \quad \text{pour tout } P \in \mathbb{Q}[X].$$

Le théorème 3.8 et la relation précédente montrent ainsi que σ est déterminé de manière unique par $\sigma(\zeta_n)$. En particulier, σ_m est l'unique élément de $\mathrm{Gal}\big(\mathbb{Q}(\zeta_n)/\mathbb{Q}\big)$ tel que $\sigma_m(\zeta_n) = \zeta_n^m$.

Par ailleurs, si Φ_n est le n-ième polynôme cyclotomique, on a

$$\Phi_n\big(\sigma(\zeta_n)\big) = \sigma\big(\Phi_n(\zeta_n)\big) = \sigma(0) = 0.$$

Ainsi, $\sigma(\zeta_n)$ est une racine de Φ_n, donc une racine primitive n-ième de l'unité. Il existe donc un entier $m \in [\![0, n-1]\!]$ premier à n tel que

$$\sigma(\zeta_n) = \zeta_n^m.$$

D'après ce qui précède, on a alors $\sigma = \sigma_m$.

Autrement dit, les éléments de $\mathrm{Gal}(\mathbb{Q}(\zeta_n)/\mathbb{Q})$ sont exactement les automorphismes σ_m, où $m \in [\![0, n-1]\!]$ est un entier premier à n. Enfin, pour tout entier $k \in \mathbb{Z}$, on a

$$\sigma_{m+kn}(\zeta_n) = \zeta_n^{m+kn} = \zeta_n^m = \sigma_m(\zeta_n).$$

On a ainsi $\sigma_{m+kn} = \sigma_m$. L'application

$$\begin{array}{rccc} \varphi : & (\mathbb{Z}/n\mathbb{Z})^\times & \longrightarrow & \mathrm{Gal}\big(\mathbb{Q}(\zeta_n)/\mathbb{Q}\big) \\ & \overline{m} & \longmapsto & \sigma_m \end{array}$$

est donc bien définie, et surjective d'après les considérations précédentes.

De plus, pour tous $\overline{m}$, $\overline{\ell} \in (\mathbb{Z}/n\mathbb{Z})^\times$, on a

$$\sigma_{\ell m}(\zeta_n) = \zeta_n^{\ell m} = (\zeta_n^m)^\ell = \sigma_\ell\big(\sigma_m(\zeta_n)\big).$$

On a donc $\sigma_{\ell m} = \sigma_\ell \sigma_m$. Cela démontre que φ est un morphisme de groupes. Il reste à établir son injectivité. Soit $m \in \mathbb{Z}$ est un entier premier à n vérifiant $\sigma_m = \mathrm{Id}$. On a donc

$$\sigma_m(\zeta_n) = \zeta_n^m = \zeta_n,$$

d'où $\zeta_n^{m-1} = 1$. Comme ζ_n est d'ordre n, on en déduit que $n \mid m-1$, soit $m \equiv 1\ [n]$. Autrement dit, $\overline{m} = \overline{1}$, et φ est donc un isomorphisme de groupes. Cela achève la démonstration. □

Revenons à notre problème de construction du p-gone régulier, où $p = 1+2^n$ est un nombre de Fermat premier. Posons $G = \mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})$. Puisque p est un nombre premier, $(\mathbb{Z}/p\mathbb{Z})^\times$ est un groupe cyclique d'ordre $p-1$ d'après le corollaire XII-2.9. D'après le lemme 4.11, il est donc de même du groupe G. Soit ρ un générateur de G.

Pour tout $i \in [\![0, n]\!]$, on pose $G_i = \langle \rho^{2^i} \rangle$ et $L_i = \mathbb{Q}(\zeta_p)^{G_i}$. Comme on a

$$G_n = \{\mathrm{Id}\} \subset G_{n-1} \subset \cdots \subset G_1 \subset G_0 = G,$$

on a

$$L_0 \subset L_1 \subset \cdots \subset L_{n-1} \subset L_n = \mathbb{Q}(\zeta_p).$$

Nous allons montrer la proposition suivante.

4.12. Proposition. *Avec les notations précédentes, on a*

(1) $L_0 = \mathbb{Q}$;

(2) *pour tout* $i \in [\![1, n]\!]$, *on a* $[L_i : L_{i-1}] = 2$. *De plus, on a* $L_i = L_{i-1}(\alpha_i)$, *avec* $\alpha_i = \sum_{k=0}^{2^{n-i}-1} \rho^{2^i k}(\zeta_p)$.

En particulier, $\mathbb{Q}(\zeta_p)/\mathbb{Q}$ *est 2-décomposable, et* ζ_p *est constructible.*

Démonstration. Avant tout, rappelons que chaque élément de G est déterminé de façon unique par l'image de ζ_p, qui est une racine primitive p-ième de l'unité. Ainsi, les éléments

$$\zeta_p,\ \rho(\zeta_p),\ \ldots,\ \rho^{p-2}(\zeta_p)$$

sont les $p-1$ puissances non triviales de ζ_p. Ils forment une famille $\mathbb{Q}$-libre, donc une $\mathbb{Q}$-base de $\mathbb{Q}(\zeta_p)$.

Soit maintenant $z \in L_0 = \mathbb{Q}(\zeta_p)^G$. Écrivons $z = a_0\zeta_p + \cdots + a_{p-2}\rho^{p-2}(\zeta_p)$, avec $a_i \in \mathbb{Q}$. Puisque $z \in L_0$, en tenant compte de l'égalité $\rho^{p-1} = \mathrm{Id}$, on a

$$z = \rho(z) = a_0\rho(\zeta_p) + \cdots + a_{p-3}\rho^{p-2}(\zeta_p) + a_{p-2}\zeta_p.$$

On en déduit les égalités $a_0 = \cdots = a_{p-2}$, d'où $z = a_0\big(\zeta_p + \cdots + \rho^{p-2}(\zeta_p)\big)$, ou encore

$$z = a_0(\zeta_p + \cdots + \zeta_p^{p-1}) = -a_0 \in \mathbb{Q}.$$

On a donc $L_0 = \mathbb{Q}$, d'où (1).

Passons au point (2). Montrons que $\alpha_i \in L_i \setminus L_{i-1}$.

On a

$$\begin{aligned}\rho^{2^i}(\alpha_i) &= \sum_{k=0}^{2^{n-i}-1} \rho^{2^i(k+1)}(\zeta_p) \\ &= \sum_{k=1}^{2^{n-i}} \rho^{2^i k}(\zeta_p) \\ &= \sum_{k=1}^{2^{n-i}-1} \rho^{2^i k}(\zeta_p) + \rho^{2^n}(\zeta_p).\end{aligned}$$

Puisque $\rho^{2^n} = \rho^{p-1} = \mathrm{Id}$, on en déduit que $\rho^{2^i}(\alpha_i) = \alpha_i$. Comme G_i est engendré par ρ^{2^i}, on en conclut que α_i est fixé par tout élément de G_i, c'est-à-dire $\alpha_i \in L_i$. Par ailleurs, on a

$$\rho^{2^{i-1}}(\alpha_i) = \sum_{k=0}^{2^{n-i}-1} \rho^{2^i k + 2^{i-1}}(\zeta_p).$$

Or, pour tout $i \in [\![1, n]\!]$, on a

$$1 \leqslant 2^i(2^{n-i} - 1) + 2^{i-1} = 2^n - 2^{i-1} = p - 1 - 2^{i-1} \leqslant p - 2.$$

Par unicité de l'écriture dans la base $\big(\zeta_p, \rho(\zeta_p), \ldots, \rho^{p-2}(\zeta_p)\big)$, on en déduit que $\rho^{2^{i-1}}(\alpha_i) \neq \alpha_i$, d'où $\alpha_i \notin L_{i-1}$.

On a ainsi $[L_i : L_{i-1}] \geqslant 2$ pour tout $i \in [\![1, n]\!]$. Par ailleurs, on a

$$[L_1 : L_0] \cdots [L_n : L_{n-1}] = [L_n : L_0] = [\mathbb{Q}(\zeta_p) : \mathbb{Q}] = \varphi(p) = p - 1 = 2^n,$$

Si un des degrés $[L_i : L_{i-1}]$ était > 2, leur produit serait $> 2^n$. On a donc

$$[L_i : L_{i-1}] = 2, \quad \text{pour tout } i \in [\![1, n]\!].$$

Comme $\alpha_i \in L_i \setminus L_{i-1}$, $(1, \alpha_i)$ est une famille L_{i-1}-libre de L_i, donc a fortiori une L_{i-1}-base de L_i. En particulier, $L_i = L_{i-1}(\alpha_i)$. Cela achève la démonstration [2]. □

En mettant bout à bout tous les résultats précédents, on a donc obtenu le résultat suivant.

4.13. Théorème (Gauss). *Soit $n \geqslant 2$ un entier. Alors, le n-gone régulier est constructible si, et seulement si, n est produit d'une puissance de 2 et de nombres de Fermat premiers deux à deux distincts.*

Les constructions du triangle équilatéral et du pentagone régulier sont classiques, et laissées au lecteur à titre d'exercice (cf. exercice 14).

2. Les arguments présentés ici sont peu ou prou ceux exposés dans l'annexe de [17].

Nous épargnerons par ailleurs au lecteur la construction du 257-gone régulier et du 65537-gone régulier, faute de place dans la marge.

Nous allons en revanche maintenant calculer explicitement la valeur exacte de $\cos\left(\frac{2\pi}{17}\right)$, découverte par Gauss à l'âge de quinze ans, ainsi que présenter une construction géométrique de l'heptadécagone régulier (i.e. le 17-gone régulier).

On vérifie que $\overline{3}$ est un générateur de $(\mathbb{Z}/17\mathbb{Z})^\times$. Le lemme 4.11 montre alors que l'unique élément $\rho \in \mathrm{Gal}(\mathbb{Q}(\zeta_{17})/\mathbb{Q})$ tel que $\rho(\zeta_{17}) = \zeta_{17}^3$ engendre $\mathrm{Gal}(\mathbb{Q}(\zeta_{17})/\mathbb{Q})$. La proposition 4.12 montre que l'on a une tour d'extensions

$$\mathbb{Q} = L_0 \subset L_1 \subset L_2 \subset L_3 \subset L_4 = \mathbb{Q}(\zeta_{17}).$$

De plus, pour tout $i \in [\![1,4]\!]$, on a $L_i = L_{i-1}(\alpha_i)$, avec $\alpha_i = \sum_{k=0}^{2^{4-i}-1} \rho^{2^i k}(\zeta_{17})$. Après calculs, on trouve

$$\begin{aligned}
\alpha_1 &= \zeta_{17} + \zeta_{17}^{-1} + \zeta_{17}^2 + \zeta_{17}^{-2} + \zeta_{17}^4 + \zeta_{17}^{-4} + \zeta_{17}^8 + \zeta_{17}^{-8} \\
\alpha_2 &= \zeta_{17} + \zeta_{17}^{-1} + \zeta_{17}^4 + \zeta_{17}^{-4} \\
\alpha_3 &= \zeta_{17} + \zeta_{17}^{-1}.
\end{aligned}$$

L'idée est maintenant de calculer le polynôme minimal de α_i sur L_{i-1}, qui est un polynôme de degré 2, afin d'expliciter successivement α_1, α_2 et α_3, sachant que $\alpha_3 = 2\cos\left(\frac{2\pi}{17}\right)$.

Pour tout $i \in [\![1,4]\!]$, posons $\beta_i = \rho^{2^{i-1}}(\alpha_i)$. On a alors

$$\rho^{2^{i-1}}(\beta_i) = \rho^{2^i}(\alpha_i) = \alpha_i,$$

puisque $\alpha_i \in L_i$. Les éléments $\alpha_i + \beta_i$ et $\alpha_i\beta_i$ sont alors des éléments de L_{i-1}. Le polynôme

$$P_i = X^2 - (\alpha_i + \beta_i)X + \alpha_i\beta_i \in L_{i-1}[X]$$

vérifie $P_i(\alpha_i) = 0$. Comme α_i est de degré 2 sur L_{i-1} et que $\mu_{\alpha_i, L_{i-1}} \mid P_i$, on a donc

$$\mu_{\alpha_i, L_{i-1}} = P_i = X^2 - (\alpha_i + \beta_i)X + \alpha_i\beta_i.$$

Un calcul montre que

$$\beta_1 = \zeta_{17}^3 + \zeta_{17}^{-3} + \zeta_{17}^5 + \zeta_{17}^{-5} + \zeta_{17}^6 + \zeta_{17}^{-6} + \zeta_{17}^7 + \zeta_{17}^{-7}.$$

On vérifie aisément que l'on a

$$\alpha_1 + \beta_1 = \sum_{k=1}^{16} \zeta_{17}^k = -1 + \sum_{k=0}^{16} \zeta_{17}^k = -1.$$

Par ailleurs, quelques calculs longuets montrent aussi que $\alpha_1\beta_1 = -4$. Ainsi, α_1 et β_1 sont les deux racines de $X^2 + X - 4$. Or, on a

$$\alpha_1 = 2\Big(\cos\left(\tfrac{2\pi}{17}\right) + \cos\left(\tfrac{4\pi}{17}\right) + \cos\left(\tfrac{8\pi}{17}\right) + \cos\left(\tfrac{16\pi}{17}\right)\Big) > 0,$$

ainsi que

$$\beta_1 = 2\Big(\cos\left(\tfrac{6\pi}{17}\right) + \cos\left(\tfrac{10\pi}{17}\right) + \cos\left(\tfrac{12\pi}{17}\right) + \cos\left(\tfrac{14\pi}{17}\right)\Big) < 0.$$

On en déduit

$$\alpha_1 = \frac{-1+\sqrt{17}}{2}, \quad \beta_1 = -\frac{1+\sqrt{17}}{2}.$$

Après quelques efforts, on obtient

$$\beta_2 = \zeta_{17}^2 + \zeta_{17}^{-2} + \zeta_{17}^8 + \zeta_{17}^{-8}, \quad \alpha_2 + \beta_2 = \alpha_1, \quad \alpha_2\beta_2 = -1.$$

Par conséquent, α_2 et β_2 sont les deux racines de $X^2 - \alpha_1 X - 1$. Un calcul approché donne $\alpha_2 > \beta_2$, et l'on obtient alors

$$\alpha_2 = \frac{1}{4}\Big(-1+\sqrt{17}+\sqrt{34-2\sqrt{17}}\Big)$$
$$\beta_2 = \frac{1}{4}\Big(-1+\sqrt{17}-\sqrt{34-2\sqrt{17}}\Big).$$

Enfin, on a

$$\beta_3 = \zeta_{17}^4 + \zeta_{17}^{-4}, \quad \alpha_3 + \beta_3 = \alpha_2, \quad \alpha_3\beta_3 = \gamma_3,$$

avec $\gamma_3 = \zeta_{17}^3 + \zeta_{17}^{-3} + \zeta_{17}^5 + \zeta_{17}^{-5} \in L_2$.

Ainsi, α_3 et β_3 sont les deux racines de $X^2 - \alpha_2 X + \gamma_3$. Il faut donc avant tout calculer explicitement γ_3.

Si $\delta_3 = \rho^2(\gamma_3)$, on a

$$\delta_3 = \zeta_{17}^6 + \zeta_{17}^{-6} + \zeta_{17}^7 + \zeta_{17}^{-7}, \quad \gamma_3 + \delta_3 = \beta_1, \quad \gamma_3\delta_3 = -1.$$

Des considérations similaires aux précédentes montrent que γ_3 et δ_3 sont les deux racines de $X^2 - \beta_1 X - 1$, puis que

$$\gamma_3 = \frac{1}{4}\Big(-1-\sqrt{17}+\sqrt{34+2\sqrt{17}}\Big)$$
$$\delta_3 = \frac{1}{4}\Big(-1-\sqrt{17}-\sqrt{34+2\sqrt{17}}\Big).$$

On peut alors maintenant calculer les deux racines de $X^2 - \alpha_2 X + \gamma_3$, qui sont respectivement $\alpha_3 = 2\cos\big(\frac{2\pi}{17}\big)$ et $\beta_3 = 2\cos\big(\frac{8\pi}{17}\big)$. Comme $\alpha_3 > \beta_3$, on obtient finalement la valeur exacte de $\cos\big(\frac{2\pi}{17}\big)$, à savoir

$$\frac{1}{16}\Big(-1+\sqrt{17}+\sqrt{34-2\sqrt{17}} + \sqrt{68+12\sqrt{17}-2(1-\sqrt{17})\sqrt{34-2\sqrt{17}}-16\sqrt{34+2\sqrt{17}}}\Big).$$

Donnons maintenant une construction géométrique explicite, due à Richmond. Nous allons suivre de près l'exposition de [8]. On conserve les notations précédentes. Soit $\varphi \in [0, \frac{\pi}{2}[$ tel que $\tan(4\varphi) = 4$. On a donc

$$X^2 + X - 4 = X^2 + 4\cot(4\varphi)X - 4.$$

Remarquons que les racines de ce polynôme sont $2\tan(2\varphi)$ et $-2\cot(2\varphi)$. En effet, on a

$$2\tan(2\varphi)\big(-2\cot(2\varphi)\big) = -4,$$

ainsi que

$$2\tan(2\varphi) - 2\cot(2\varphi) = -4\frac{\cos^2(2\varphi) - \sin^2(2\varphi)}{2\sin(2\varphi)\cos(2\varphi)} = -4\cot(4\varphi).$$

Ainsi, comme α_1 et β_1 sont racines de ce polynôme et que $\alpha_1 > 0$, on a

$$\alpha_1 = 2\tan(2\varphi), \quad \beta_1 = -2\cot(2\varphi).$$

De plus, α_2 et β_2 sont racines de $X^2 - \alpha_1 X - 1 = X^2 - 2\tan(2\varphi)X - 1$. Or, les racines de ce polynômes sont $\tan\left(\varphi + \frac{\pi}{4}\right)$ et $\tan\left(\varphi - \frac{\pi}{4}\right)$. En effet, on a

$$\tan\left(\varphi + \tfrac{\pi}{4}\right) = \frac{\tan(\varphi) + \tan\left(\frac{\pi}{4}\right)}{1 - \tan(\varphi)\tan\left(\frac{\pi}{4}\right)} = \frac{\tan(\varphi) + 1}{1 - \tan(\varphi)},$$

ainsi que

$$\tan\left(\varphi - \tfrac{\pi}{4}\right) = \frac{\tan(\varphi) - \tan\left(\frac{\pi}{4}\right)}{1 + \tan(\varphi)\tan\left(\frac{\pi}{4}\right)} = \frac{\tan(\varphi) - 1}{1 + \tan(\varphi)},$$

On a donc $\tan\left(\varphi + \frac{\pi}{4}\right)\tan\left(\varphi - \frac{\pi}{4}\right) = -1$. De plus, on obtient aussi

$$\tan(2\varphi) = \tan\left(\left(\varphi + \tfrac{\pi}{4}\right) + \left(\varphi - \tfrac{\pi}{4}\right)\right) = \frac{\tan\left(\varphi + \frac{\pi}{4}\right) + \tan\left(\varphi - \frac{\pi}{4}\right)}{1 - \tan\left(\varphi + \frac{\pi}{4}\right)\tan\left(\varphi - \frac{\pi}{4}\right)},$$

soit

$$\tan(2\varphi) = \frac{\tan\left(\varphi + \frac{\pi}{4}\right) + \tan\left(\varphi - \frac{\pi}{4}\right)}{2},$$

d'où $\tan\left(\varphi + \frac{\pi}{4}\right) + \tan\left(\varphi - \frac{\pi}{4}\right) = 2\tan(2\varphi)$.

Puisque $\alpha_2 > \beta_2$, on a donc

$$\alpha_2 = \tan\left(\varphi + \tfrac{\pi}{4}\right), \quad \beta_2 = \tan\left(\varphi - \tfrac{\pi}{4}\right).$$

Comme γ_3 et δ_3 sont racines de $X^2 - \beta_1 X - 1 = X^2 + 2\cot(2\varphi)X - 1$, et que $\gamma_3 > 0$, on a

$$\gamma_3 = \tan(\varphi), \quad \delta_3 = -\cot(\varphi).$$

En effet, on a

$$\tan(\varphi)\big(-\cot(\varphi)\big) = -1,$$

et un calcul similaire à celui au début de ces considérations montre que

$$\tan(\varphi) - \cot(\varphi) = -2\cot(2\varphi).$$

Remarquons enfin que l'on a

$$(\zeta_{17}^3 + \zeta_{17}^{-3}) + (\zeta_{17}^5 + \zeta_{17}^{-5}) = \gamma_3,$$

et

$$(\zeta_{17}^3 + \zeta_{17}^{-3})(\zeta_{17}^5 + \zeta_{17}^{-5}) = \zeta_{17}^8 + \zeta_{17}^{-2} + \zeta_{17}^2 + \zeta_{17}^{-8} = \beta_2.$$

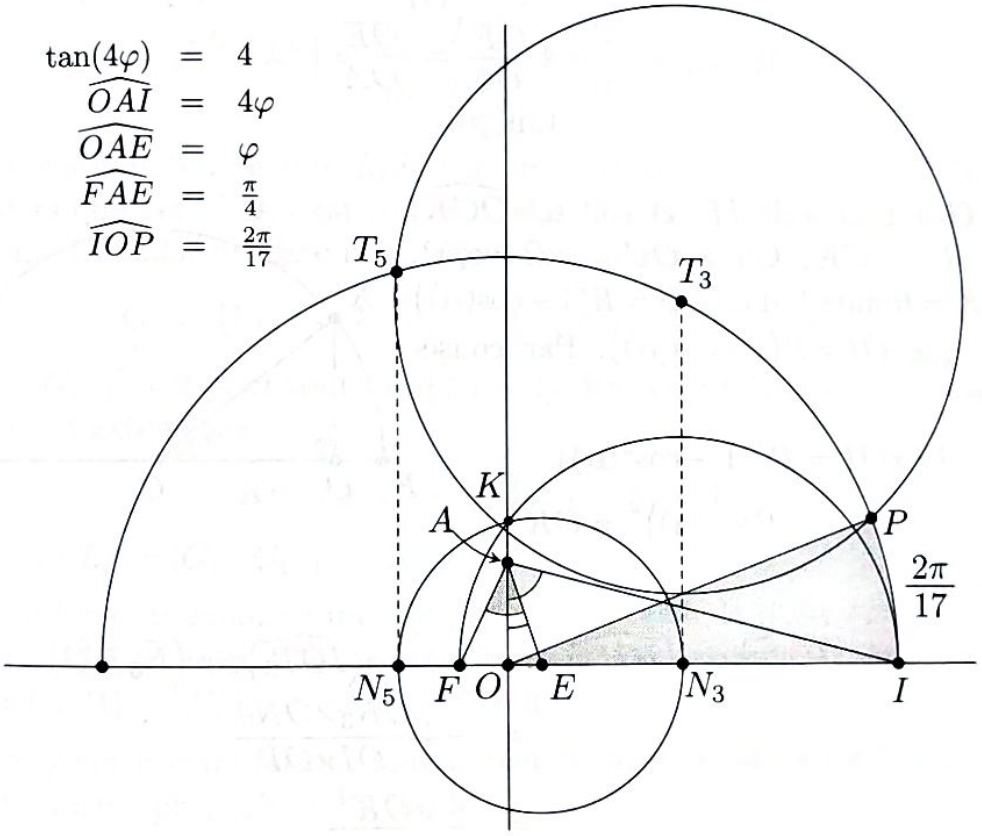
Par conséquent, si $\theta = \frac{2\pi}{17}$, on a

$$2\big(\cos(3\theta) + \cos(5\theta)\big) = \tan(\varphi) \quad \text{et} \quad 4\cos(3\theta)\cos(5\theta) = \tan\left(\varphi - \tfrac{\pi}{4}\right).$$

On construit le point $A = \left(0, \frac{1}{4}\right)$, si bien que $\tan\left(\widehat{OAI}\right) = 4$. En utilisant deux bissections successives, on construit le point E sur l'axe des abscisses tel que l'angle $\widehat{OAE} = \frac{1}{4}\widehat{OAI}$. On a donc

$$\tan\left(4\widehat{OAE}\right) = \tan\left(\widehat{OAI}\right) = 4.$$

On a ainsi $\widehat{OAI} = 4\varphi$ et $\widehat{OAE} = \varphi$.



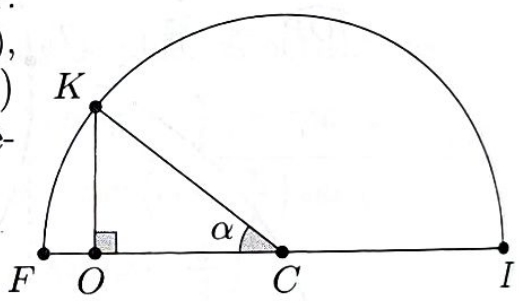
En construisant la perpendiculaire à (AE) passant par A, puis à l'aide d'une bissection, on obtient l'unique point F sur l'axe des abscisses tel que l'angle $\widehat{FAE}$ soit de mesure $\frac{\pi}{4}$.

Le demi-cercle supérieur de diamètre IF coupe l'axe des ordonnées en un point K. On construit enfin le cercle de centre E et de rayon EK. Celui-ci coupe la droite (OI) en deux points N_3 et N_5, projections orthogonales respectives des points T_3 et T_5 du demi-cercle-unité supérieur, N_3 étant celui appartenant au segment OI.

On a alors

$$\begin{aligned}
2\Big(\cos\big(\widehat{IOT_3}\big)+\cos\big(\widehat{IOT_5}\big)\Big) &= 2\Big(\cos\big(\widehat{IOT_3}\big)+\cos\big(\widehat{N_5T_5O}+\tfrac{\pi}{2}\big)\Big)\\
&= 2\Big(\cos\big(\widehat{IOT_3}\big)-\sin\big(\widehat{N_5T_5O}\big)\Big)\\
&= 2\Big(\frac{ON_3}{OT_3}-\frac{ON_5}{OT_5}\Big) = 2\Big(\frac{ON_3}{OI}-\frac{ON_5}{OI}\Big)\\
&= 2\Big(\frac{OE+EN_3}{OI}-\frac{EN_5-OE}{OI}v\Big)\\
&= 2\Big(\frac{OE+EN_3}{OI}-\frac{EN_3-OE}{OI}\Big)\\
&= 4\,\frac{OE}{OI} = \frac{OE}{OA}\\
&= \tan(\varphi).
\end{aligned}$$

Soit C le milieu de IF, et soit $\alpha=\widehat{OCK}$. Soit $R = CK$. On a $OC = R\cos(\alpha)$, et $OK=R\sin(\alpha)$, d'où $OF=R\big(1-\cos(\alpha)\big)$ ainsi que $OI=R\big(1+\cos(\alpha)\big)$. Par conséquent,



$$\begin{aligned}
OF\times OI &= R^2\big(1-\cos^2(\alpha)\big)\\
&= \big(R\sin(\alpha)\big)^2 = OK^2.
\end{aligned}$$

Mais alors,

$$\begin{aligned}
4\cos\big(\widehat{IOT_3}\big)\cos\big(\widehat{IOT_5}\big) &= -4\cos\big(\widehat{IOT_3}\big)\sin\big(\widehat{N_5T_5O}\big)\\
&= -4\,\frac{ON_3\times ON_5}{OI\times OI}\\
&= -4\,\frac{OK^2}{OI^2}\\
&= -4\,\frac{OF\times OI}{OI^2}\\
&= -4\,\frac{OF}{OI}\\
&= -\frac{OF}{OA}\\
&= \tan\big(\varphi-\tfrac{\pi}{4}\big).
\end{aligned}$$

On a donc $\widehat{IOT_3} = 3\theta$ et $\widehat{IOT_5} = 5\theta$. Le demi-cercle-unité supérieur et le cercle de centre T_3 et de rayon T_3T_5 se coupent en un point $P \neq T_5$ du cercle-unité tel que $\widehat{POT_3} = 2\theta$. Mais alors, $\widehat{IOP} = \theta$, et l'on obtient finalement $P = \big(\cos\big(\frac{2\pi}{17}\big), \sin\big(\frac{2\pi}{17}\big)\big)$.

5. Et après...

La démonstration du théorème de Gauss peut sembler un peu miraculeuse. En fait, elle ne l'est pas tant que ça.

Si L est un corps et G est un sous-groupe fini du groupe des automorphismes d'anneaux de L, nous démontrerons dans un chapitre ultérieur que l'extension L/L^G est de degré fini et que

$$[L : L^G] = |G|.$$

En particulier, si H est un sous-groupe de G, on a $L^G \subset L^H$, et

$$[L^H : L^G] = \frac{[L : L^G]}{[L : L^H]} = \frac{|G|}{|H|} = [G : H].$$

Supposons maintenant que L/K soit une extension de degré fini. On vérifie alors que $\mathrm{Gal}(L/K)$ est fini. Si $\mathrm{Gal}(L/K)$ est un 2-groupe, d'après le théorème IX-3.23, il existe une suite de composition

$$G_r = \{\mathrm{Id}_L\} \lhd G_{r-1} \lhd \cdots \lhd G_1 \lhd G_0 = \mathrm{Gal}(L/K),$$

où $[G_i : G_{i+1}] = 2$ pour tout $i \in [\![0, r-1]\!]$. En posant $L_i = L^{G_i}$, on obtient une tour d'extensions

$$K \subset L_0 \subset \cdots \subset L_{r-1} \subset L_r = L,$$

où $[L_{i+1} : L_i] = [G_i : G_{i+1}] = 2$.

Cela explique de manière un peu plus conceptuelle le point (2) de la proposition 4.12, puisque si p est un nombre de Fermat premier, le groupe de Galois $\mathrm{Gal}\big(\mathbb{Q}(\zeta_p)/\mathbb{Q}\big)$ est cyclique d'ordre $p-1$, donc un 2-groupe fini.

L'autre point crucial faisant fonctionner la démonstration est le point (1) de cette même proposition, à savoir

$$\mathbb{Q}(\zeta_p)^{\mathrm{Gal}(\mathbb{Q}(\zeta_p)/\mathbb{Q})} = \mathbb{Q}.$$

On peut donc se poser légitimement la question suivante : soit L/K une extension de degré fini. A-t-on $L^{\mathrm{Gal}(L/K)} = K$?

La réponse est négative, comme le montrent les exemples suivants.

5.1. Exemples

(1) Soit $K = \mathbb{Q}$, soit $\alpha = \sqrt[3]{2}$, et soit $L = \mathbb{Q}(\alpha)$.
Si $\sigma \in \mathrm{Gal}(L/K)$, on a

$$\sigma(\alpha)^3 = \sigma(\alpha^3) = \sigma(2) = 2.$$

Comme $L \subset \mathbb{R}$, $\sigma(\alpha)$ est une racine réelle du polynôme $\mu_{\alpha,K} = X^3 - 2$. On a donc $\sigma(\alpha) = \alpha$.

Comme $(1, \alpha, \alpha^2)$ est une K-base de L, on en déduit que $\sigma = \mathrm{Id}_L$. Par conséquent, $\mathrm{Gal}(L/K) = \{\mathrm{Id}_L\}$, et ainsi

$$L^{\mathrm{Gal}(L/K)} = L \neq K.$$

(2) Soit $K = \mathbb{F}_2(T^2)$. Le polynôme $f = X^2 - T^2 \in K[X]$ est irréductible, puisque $T \notin \mathbb{F}_2(T^2)$, comme le lecteur le vérifiera à titre d'exercice. Soit $\alpha = T$, et $L = K(\alpha)$. On a alors

$$\mu_{\alpha,K} = X^2 - T^2 = (X - T)^2,$$

puisque K est de caractéristique 2. Un raisonnement similaire au précédent montre alors que $\mathrm{Gal}(L/K) = \{\mathrm{Id}_L\}$, et donc que

$$L^{\mathrm{Gal}(L/K)} = L \neq K.$$

Un exemple d'extension $L/\mathbb{Q}$ de degré 4 et telle que $L^{\mathrm{Gal}(L/\mathbb{Q})} \neq \mathbb{Q}$ est donné dans l'exercice 15.

Analysons un peu les deux contre-exemples précédents. Dans chacun des cas, le polynôme minimal de α sur K n'a pas assez de racines dans L, et cela pour deux raisons distinctes. Dans le premier cas, $\mu_{\alpha,K}$ a toutes ses racines distinctes (dans $\mathbb{C}$), mais une seule appartient à L. Dans le second cas, c'est parce que le polynôme minimal de α sur K possède une racine multiple.

Les extensions L/K de degré fini vérifiant $L^{\mathrm{Gal}(L/K)} = K$ jouissent en fait de propriétés particulières : ce sont les extensions dites *galoisiennes*. Elles seront définies et étudiées dans un chapitre ultérieur. Pour l'instant, disons simplement que ce sont exactement les extensions pour lesquelles les deux phénomènes précédents ne peuvent se produire.

Ces exemples montrent également que tous les sous-corps d'une extension de corps L/K ne sont pas forcément de la forme L^H, où H est un sous-groupe de $\mathrm{Gal}(L/K)$. Nous verrons néanmoins que c'est le cas lorsque L/K est galoisienne.

Concluons ce chapitre en dégageant un point important qui revient constamment dans les énoncés ou les démonstrations, à savoir qu'il est souvent utile de se placer dans une extension de K dans laquelle un polynôme de $K[X]$ donné possède au moins une racine (voire se scinde), par exemple pour construire de nouveaux corps, ou pour calculer le groupe de Galois d'une extension donnée. Lorsque K est un sous-corps de $\mathbb{C}$, il n'y a aucun problème, puisque tout polynôme se scinde dans $\mathbb{C}[X]$. En revanche, la question est plus délicate lorsque K est un corps quelconque. Cette question est presque résolue (au moins dans le cas d'un polynôme irréductible) par le lemme XXI-1.10, modulo un tour de passe-passe ensembliste que nous expliquerons dans le chapitre suivant.

En revanche, il reste un point délicat à soulever : la notion d'élément algébrique est a priori relative à une extension L/K ambiante. Or, moralement, ce qui définit vraiment les propriétés d'un élément algébrique, c'est son polynôme minimal, qui est à coefficients dans K. Il serait donc bon de disposer d'une extension de référence dans laquelle tout polynôme de $K[X]$ a une racine. Au vu ce ce qui précède, il serait même extrêmement utile d'avoir l'existence d'une extension E/K jouissant de propriétés similaires à celles de $\mathbb{C}$, à savoir que tout polynôme de $E[X]$ est scindé. Cela conduit à la notion de corps algébriquement clos. Le chapitre suivant sera consacré à la résolution de toutes ces questions.

Nous reviendrons ensuite sur l'étude des morphismes d'extensions, en particulier des automorphismes d'une extension L/K de degré fini. Enfin, nous étudierons les extensions galoisiennes de manière approfondie.

6. Exercices

E1. Construire une extension $L/\mathbb{Q}$ algébrique de degré infini.

E2. Montrer les égalités suivantes :

(i) $\mathbb{Q}\left(\frac{-1+\sqrt{5}}{4}\right) = \mathbb{Q}(\sqrt{5})$;

(ii) $\mathbb{Q}(\sqrt{12}, i\sqrt{3}, \sqrt[3]{27}) = \mathbb{Q}(i, \sqrt{3})$;

(iii) $\mathbb{Q}(\sqrt{2}+\sqrt{3}) = \mathbb{Q}(\sqrt{2}, \sqrt{3})$.

E3. Dans chacun des cas suivants, montrer que α est algébrique sur K, et calculer son polynôme minimal sur K.

(i) $K = \mathbb{Q},\ \alpha = \sqrt{6}$;

(ii) $K = \mathbb{Q},\ \alpha = 1 + \sqrt[3]{7}$;

(iii) $K = \mathbb{Q}(\sqrt{3}),\ \alpha = \sqrt{2}$.

E4. Soit $P = X^3 + 2X + 2 \in \mathbb{Q}[X]$, et soit α une racine de P dans $\mathbb{C}$.

a. Montrer que P est irréductible dans $\mathbb{Q}[X]$. Que vaut $[\mathbb{Q}(\alpha) : \mathbb{Q}]$?

b. Exprimer $u = \alpha^{-1}$, $v = \alpha^6 + \alpha^4 + 3\alpha^3 - \alpha^2 + 3$ et $w = (\alpha^2 + \alpha + 1)^{-1}$ en fonction de 1, α et α^2.

c. Quel est le polynôme minimal de v sur $\mathbb{Q}$?

E5. Expliciter l'inverse de $\sqrt[3]{2} - 1 \in \mathbb{Q}(\sqrt[3]{2})$ dans la base $\left(1, \sqrt[3]{2}, (\sqrt[3]{2})^2\right)$.

E6.

a. Montrer que i et $j = -\frac{1}{2} + i\frac{\sqrt{3}}{2}$ sont algébriques sur $\mathbb{Q}$, et déterminer leurs degrés sur $\mathbb{Q}$.

b. Calculer $[\mathbb{Q}(\sqrt{3},i):\mathbb{Q}]$, $[\mathbb{Q}(\sqrt{3},j):\mathbb{Q}]$ et $[\mathbb{Q}(\sqrt{3},i,j):\mathbb{Q}]$.

c. Comparer $\mathbb{Q}(\sqrt{3},i)$ et $\mathbb{Q}(\sqrt{3}+i)$.

d. Déterminer le polynôme minimal de $\sqrt{3}+i$ sur $\mathbb{Q}$.

E7.

a. Déterminer $[\mathbb{Q}(\sqrt[3]{3},\sqrt{5}):\mathbb{Q}]$, et donner une $\mathbb{Q}$-base de $\mathbb{Q}(\sqrt[3]{3},\sqrt{5})$.

b. Comparer $[\mathbb{Q}(\sqrt[3]{3},\sqrt{5}):\mathbb{Q}]$ et $[\mathbb{Q}(\sqrt[3]{3}+\sqrt{5}):\mathbb{Q}]$.

Indication. À l'aide de la base calculée précédemment, montrer qu'un polynôme $P\in\mathbb{Q}[X]$ de degré $\leqslant 3$ annulant $\sqrt[3]{3}+\sqrt{5}$ est nécessairement nul.

c. Déterminer le polynôme minimal de $\sqrt[3]{3}+\sqrt{5}$ sur $\mathbb{Q}$.

E8. Montrer que l'ensemble des nombres complexes algébriques sur $\mathbb{Q}$ est dénombrable.

E9. Soit L/K une extension de degré impair. Montrer que si $a\in K$ est un carré dans L, alors c'est un carré dans K.

E10. Soit L/K une extension, et soit $\alpha\in L$ tel que $K(\alpha)/K$ est de degré impair. Montrer que $K(\alpha)=K(\alpha^2)$. Peut-on se passer de l'hypothèse d'imparité ?

E11. Soit L/K une extension de degré premier. Décrire toutes les sous-extensions de L/K.

E12. Soit L/K une extension, et soient E/K et E'/K deux sous-extensions de L/K. Si $[E':K]=2$, montrer que E/K et E'/K sont linéairement disjointes si, et seulement si, $E\cap E'=K$.

E13. D'après l'exercice 8, l'ensemble des nombres transcendants a la puissance du continu. On veut exhiber ici une famille indénombrable de nombres transcendants.

On rappelle que tout $x\in\,]0,1]$ s'écrit de manière unique sous la forme

$$x=\sum_{k\geqslant 1}\frac{d_k}{10^k},$$

où $(d_k)_{k\geqslant 1}$ est une suite d'entiers de $[\![0,9]\!]$ ne convergeant pas[3] vers 0.

3. Autrement dit, lorsque x est un nombre décimal, on privilégie son développement impropre, plutôt que son développement propre.

On appelle *nombre de Liouville* tout réel de $]0,1]$ de la forme

$$\sum_{k\geqslant 1} \frac{a_k}{10^{k!}},$$

où $(a_k)_{k\geqslant 1}$ est une suite d'entiers de $[\![0,9]\!]$ ne convergeant pas vers 0.

a. Montrer que l'ensemble des nombres de Liouville n'est pas dénombrable.

b. Soit $x \in \mathbb{R}$ un nombre réel algébrique sur $\mathbb{Q}$ de degré $d \geqslant 1$. Montrer qu'il existe une constante $M > 0$ telle que pour tous $a, b \in \mathbb{Z}$, $b > 1$ tels que $x \neq \frac{a}{b}$, on a

$$\left|x - \frac{a}{b}\right| \geqslant \frac{M}{b^d}.$$

Indication. On distinguera les cas $d = 1$ et $d > 1$. Si $d > 1$, considérer un polynôme $P \in \mathbb{Z}[X]$ tel que $P(x) = 0$. Montrer que $\left|P(\frac{a}{b})\right| \geqslant \frac{1}{b^d}$, et utiliser le théorème des accroissements finis sur $[x-1, x+1]$.

c. Soit $(a_k)_{k\geqslant 1}$ une suite d'entiers de $[\![0,9]\!]$ ne convergeant pas vers 0, et soit $x = \sum_{k\geqslant 1} \frac{a_k}{10^{k!}}$. Pour tout $n \geqslant 1$, on pose

$$x_n = \sum_{k=1}^{n} \frac{a_k}{10^{k!}}.$$

Montrer que pour tout $n \geqslant 1$, on a $|x - x_n| \leqslant \frac{1}{10^{n\times n!}}$.

d. Déduire des questions précédentes que tout nombre de Liouville est transcendant.

E14.

a. Proposer une construction du triangle équilatéral.

b. Déterminer la valeur exacte de $\cos\left(\frac{2\pi}{5}\right)$, et en déduire une construction du pentagone régulier à la règle et au compas.

Indication. On pourra s'inspirer de la méthode de calcul de $\cos\left(\frac{2\pi}{17}\right)$, ou bien remarquer que

$$\zeta_5^2 + \zeta_5^{-2} + \zeta_5 + \zeta_5^{-1} + 1 = 0.$$

E15. Le but de cet exercice est de construire une famille infinie de nombres non constructibles de degré 4 sur $\mathbb{Q}$. Soit p un nombre premier.

a. Montrer que $P(X) = X^4 + pX - p$ et $Q(X) = X^3 + 4pX - p^2$ sont irréductibles dans $\mathbb{Q}[X]$.

b. Montrer que le polynôme P a exactement deux racines réelles x_1, x_2. En déduire $[\mathbb{Q}(x_1) : \mathbb{Q}]$ et $[\mathbb{Q}(x_2) : \mathbb{Q}]$.

On pose $a = -(x_1 + x_2)$, $b = x_1 x_2$ (donc $a, b \in \mathbb{R}$).

c. Montrer l'existence de $a', b' \in \mathbb{R}$ tels que

$$P(X) = (X^2 + aX + b)(X^2 + a'X + b').$$

Montrer les relations

$$a' = -a, \quad b + b' = a^2, \quad a(b' - b) = p, \quad bb' = -p.$$

d. En déduire que b et b' sont les racines de $X^2 - a^2 X - p \in \mathbb{R}[X]$.

e. En déduire que $b' - b = \pm\sqrt{a^4 + 4p}$ et que a^2 est une racine de Q.

f. Calculer $[\mathbb{Q}(a^2) : \mathbb{Q}]$, et en déduire que a^2 et a ne sont pas constructibles.

g. Montrer qu'au moins un des réels x_1, x_2 n'est pas constructible.

h. En déduire l'existence d'une extension de corps $K/\mathbb{Q}$ de degré 4 qui n'est pas 2-décomposable.

i. Montrer que $x_2 \notin \mathbb{Q}(x_1)$.

Indication. Montrer que si $x_2 \in \mathbb{Q}(x_1)$, alors $a^2 \in \mathbb{Q}(x_1)$.

j. En déduire que le seul automorphisme de $\mathbb{Q}$-algèbres de $\mathbb{Q}(x_1)$ est le morphisme identité.

E16. Discuter la quintisection de l'angle (division en cinq angles égaux) à la règle et au compas.

Chapitre XXVI

Polynômes et racines

Dans ce chapitre, nous nous intéressons à certaines des questions soulevées à la fin du chapitre précédent, à savoir la construction d'une extension de corps dans laquelle un polynôme donné possède une racine (voire se scinde en produit de facteurs de degré 1).

1. Existence de racines dans une extension et applications

Il est toujours plus agréable de faire des calculs avec les polynômes lorsque l'on dispose de racines (on peut alors par exemple utiliser les relations entre coefficients et racines). Malheureusement, il arrive souvent qu'un polynôme de $K[X]$ n'ait pas de racine dans K, par exemple s'il est irréductible. On va maintenant voir que l'on peut construire une extension de corps L/K dans laquelle un polynôme de $K[X]$ donné se décompose totalement en facteurs de degré 1.

Introduisons quelques notations. Si $\iota : A \longrightarrow B$ est un morphisme d'anneaux commutatifs, alors il se prolonge en un morphisme d'anneaux

$$\begin{array}{rccc} \iota : & A[X] & \longrightarrow & B[X] \\ & \displaystyle\sum_{m\geqslant 0} a_m X^m & \longmapsto & \displaystyle\sum_{m\geqslant 0} \iota(a_m) X^m. \end{array}$$

De plus, ce morphisme est injectif/surjectif/bijectif dès que le morphisme précédent l'est.

Remarquons aussi que ι est compatible avec l'évaluation dans le sens suivant : pour tout $f \in A[X]$ et tout $\alpha \in A$, on a

$$\iota(f)\big(\iota(\alpha)\big) = \iota\big(f(\alpha)\big).$$

Nous allons commencer ce paragraphe en expliquant comment construire une extension L/K dans laquelle un polynôme $f \in K[X]$ non constant a une racine. Bien entendu, il suffit de savoir traiter le cas où f est irréductible. Au vu du lemme XXI-1.10, l'idée naturelle est alors de considérer l'anneau quotient $K[X]/(f)$, mais comme on l'a déjà remarqué, ce corps ne convient pas tout à fait. Pour y remédier, nous avons besoin d'un tour de passe-passe ensembliste très classique, appelé transport de structure.

1.1. Lemme (Transport de structure). *Soient K et M deux corps, et soit $\iota : K \longrightarrow M$ un morphisme d'anneaux (nécessairement injectif). Alors, il existe un corps L contenant K comme sous-corps, dont les opérations prolongent celles de K et un isomorphisme de K-algèbres $\sigma : L \xrightarrow{\sim} M$ prolongeant ι.*

En particulier, pour tout $f \in K[X]$, f possède une racine dans L si, et seulement si, $\iota(f) \in M[X]$ possède une racine dans M.

Démonstration. Considérons $M' = M \setminus \iota(K)$. Quitte à remplacer M' par un ensemble équipotent à M' adéquat, on peut supposer que K et M' sont disjoints. On pose alors $L = K \cup M'$. Par construction, $K \subset L$. On définit une application $\sigma : L \longrightarrow M$ par

$$\sigma_{|K} = \iota \quad \text{et} \quad \sigma_{|M'} = \mathrm{Id}_{M'}.$$

Autrement dit,

$$\sigma(x) = \iota(x), \ \text{ si } x \in K \quad \text{et} \quad \sigma(x) = x, \ \text{ si } x \in M'.$$

C'est une bijection ensembliste. On définit alors une somme et un produit sur L par

$$x + y = \sigma^{-1}\big(\sigma(x) + \sigma(y)\big) \quad \text{et} \quad xy = \sigma^{-1}\big(\sigma(x)\sigma(y)\big).$$

On vérifie aisément que L est un corps pour les opérations ci-dessus, et que cette structure de corps prolonge celle de K (i.e. la restriction de ces opérations à K sont les opérations de K). Ainsi, L/K est une extension de corps. De plus, σ est alors un morphisme de K-algèbres. Ces vérifications sont laissées en exercice au lecteur.

Soit $f \in K[X]$. Pour tout $\alpha \in L$, on a

$$f(\alpha) = 0 \iff \sigma\big(f(\alpha)\big) = 0 \iff \sigma(f)\big(\sigma(\alpha)\big) = 0 \iff \iota(f)\big(\sigma(\alpha)\big) = 0,$$

la première équivalence découlant du fait que σ est bijective, et la dernière du fait que $\sigma_{|K} = \iota$. La dernière partie en découle puisque $M = \sigma(L)$. Cela achève la démonstration. □

On peut maintenant démontrer le résultat attendu.

1.2. Proposition. *Soit K un corps, et soit $f \in K[X]$ non constant. Alors, il existe une extension L/K dans laquelle f admet au moins une racine.*

Démonstration. Si f admet une racine dans K, on pose $L = K$. Sinon, f possède un facteur irréductible g de degré $\geqslant 2$. Posons $M = K[X]/(g)$, et soit

$$\begin{aligned} \iota : K &\longrightarrow M \\ \lambda &\longmapsto \overline{\lambda}. \end{aligned}$$

D'après le lemme XXI-1.10, M est un corps dans lequel $\iota(g)$ admet une racine. Comme $g \mid f$, on a $\iota(g) \mid \iota(f)$ et $\iota(f)$ possède également une racine dans M. On conclut alors par transport de structure. □

1.3. Remarque. Les difficultés techniques rencontrées ici montrent que l'inclusion de corps n'est pas le bon cadre pour travailler. En fait, il faut généraliser la notion d'extension en disant que c'est un couple (L, ι), où L est un corps et $\iota : K \longrightarrow L$ est un morphisme d'anneaux (nécessairement injectif).

Comme application de l'existence d'une telle extension, on continue par un critère d'irréductibilité d'un polynôme en termes de non-existence de racines, et qui généralise à la fois les critères connus en degré 2 et 3, et la proposition XXI-6.14.

1.4. Proposition. *Soit K un corps, et soit $f \in K[X]$ de degré $n \geqslant 1$. Alors, le polynôme f est irréductible si, et seulement si, pour toute extension L/K de degré $\leqslant \frac{n}{2}$, f n'a pas de racines dans L.*

Démonstration. Si f n'est pas irréductible, alors $f = gh$ où $g, h \in K[X]$ sont non constants. Au moins un de ces deux facteurs est de degré $\leqslant \frac{n}{2}$ (sinon on aurait $\deg(gh) \geqslant n + 1 = \deg(f)$), disons g. Soit π un facteur irréductible de g, qui est donc aussi de degré $\leqslant \frac{n}{2}$, et soit E/K une extension dans laquelle π a une racine $\alpha \in E$. Par définition, on a $\mu_{\alpha,K} = \pi$. Ainsi, $K(\alpha)/K$ est une extension de degré $\leqslant \frac{n}{2}$ dans lequel π a une racine, et par suite f possède une racine dans $K(\alpha)$ (car $\pi \mid g \mid f$).

Inversement, si f a une racine $\alpha \in L$, avec $[L : K] \leqslant \frac{n}{2}$, soit $g = \mu_{\alpha,K}$ le polynôme minimal de α sur K. Alors, puisque $f(\alpha) = 0$, on a $g \mid f$, donc f n'est pas irréductible puisque $1 \leqslant \deg(g) < n$. □

1.5. Remarque. Soit $K = \mathbb{F}_p$. Puisque $\mathbb{F}_p$ n'a qu'une seule extension de degré d donné, à savoir $\mathbb{F}_{p^d}/\mathbb{F}_p$, par unicité d'un corps à p^d éléments, on retrouve alors le résultat de la proposition XXI-6.14,

On a alors le résultat suivant.

1.6. Proposition. *Soit K un corps, et soit $f \in K[X]$ non constant. Alors, il existe une extension L/K dans laquelle f est scindé (i.e. se décompose en facteurs de degré 1).*

Démonstration. On raisonne par récurrence sur le degré de f. Plus précisément, l'hypothèse de récurrence est la suivante :

(H_n) : le résultat est vrai pour toute extension de corps M/K et tout polynôme $f \in M[X]$ de degré $n \geqslant 1$.

Si $n = 1$, on prend $L = M$. Supposons avoir démontré le résultat pour tous les polynômes de degré n à coefficients dans une extension de K et soit f un polynôme de degré $n+1$ à coefficients dans une extension M/K. Si f se décompose en facteurs de degré 1, on prend $L = M$. Sinon, f admet un facteur irréductible g de degré $\geqslant 2$. D'après la proposition précédente, il existe une extension M'/M dans laquelle g, et donc a fortiori f, a une racine $\alpha \in M'$. Mais alors, on a $f = (X - \alpha)h$, $h \in M'[X]$. Comme h est de degré n, par hypothèse de récurrence, il existe une extension L/M' dans laquelle h se décompose en facteurs de degré 1. Mais alors, f est aussi produit de facteurs de degré 1 dans L. Cela achève la récurrence.

Il suffit alors de prendre $M = K$ pour conclure. □

1.7. Remarque. Si L/K est une extension de corps dans laquelle un polynôme $f \in K[X]$ de degré $n \geqslant 1$ est scindé, on a

$$f = c(X - \alpha_1) \cdots (X - \alpha_n), \quad c \in L^\times, \ \alpha_i \in L.$$

Ainsi, f a n racines (non nécessairement distinctes) dans L.

Ce simple résultat permet déjà d'avoir des résultats intéressants.

1.8. Proposition. *Soit K un corps, et soit*

$$P = a_n X^n + a_{n-1} X^{n-1} + \cdots + a_1 X + a_0 \in K[X], \quad \deg(P) = n \geqslant 1.$$

Soient $\alpha_1, \cdots, \alpha_n$ les racines de P dans une extension L/K dans laquelle P est scindé. Alors, on a

$$\sigma_j(\alpha_1, \cdots, \alpha_n) = (-1)^j \frac{a_{n-j}}{a_n}, \quad \text{pour tout } j \in [\![1, n]\!],$$

où $\sigma_1, \ldots, \sigma_n \in K[X_1, \ldots, X_n]$ *sont les polynômes symétriques élémentaires.*

Démonstration. On écrit $P = a_n(X-\alpha_1)\cdots(X-\alpha_n)$, et l'on développe. □

1.9. Théorème. *Soit K un corps quelconque, soit $P \in K[X]$ un polynôme de degré $n \geqslant 1$, et soient $\alpha_1, \ldots, \alpha_n$ les racines de P dans une extension L/K dans laquelle P est scindé.*

Alors, pour tout polynôme **symétrique** *$f \in K[X_1, \ldots, X_n]$, on a*

$$f(\alpha_1, \ldots, \alpha_n) \in K,$$

et cette valeur est indépendante du choix de L.

Autrement dit, toute expression polynomiale symétrique en les racines de P est un élément de K, indépendant du choix de L.

Plus précisément[1]*, si $P = a_nX^n + \cdots + a_1X + a_0$, si f est de degré partiel égal à d, et si A est un sous-anneau de K contenant les coefficients de f, il existe un polynôme $h \in A[T_0, \ldots, T_{n-1}, T_n]$, indépendant du choix de L, tel que*

$$f(\alpha_1, \ldots, \alpha_n) = \frac{h(a_0, \ldots, a_n)}{a_n^d}.$$

En particulier, si P est unitaire, toute expression polynomiale symétrique en les racines de P est un polynôme en les coefficients de P.

Démonstration. Démontrons la première partie. On suppose que f est symétrique, de degré partiel égal à d. Par le théorème de structure des polynômes symétriques, on peut écrire

$$f(X_1, \ldots, X_n) = Q\big(\sigma_1(X_1, \ldots, X_n), \ldots, \sigma_n(X_1, \ldots, X_n)\big),$$

où $Q \in A[T_1, \ldots, T_n]$ est un polynôme de degré d. On a alors

$$\begin{aligned} f(\alpha_1, \ldots, \alpha_n) &= Q\big(\sigma_1(\alpha_1, \ldots, \alpha_n), \ldots, \sigma_n(\alpha_1, \ldots, \alpha_n)\big) \\ &= Q\Big(-\frac{a_{n-1}}{a_n}, \ldots, (-1)^n \frac{a_0}{a_n}\Big). \end{aligned}$$

Écrivons

$$Q = \sum_{m_1+\cdots+m_n \leqslant d} c_{m_1,\ldots,m_n} T_1^{m_1} \cdots T_n^{m_n}, \quad c_{m_1,\ldots,m_n} \in A.$$

On a alors

$$f(\alpha_1, \ldots, \alpha_n) = \sum_{m_1+\cdots+m_n \leqslant d} c_{m_1,\ldots,m_n} (-1)^{m_1+2m_2+\cdots+nm_n} \frac{a_{n-1}^{m_1} \cdots a_0^{m_n}}{a_n^{m_1+\cdots+m_n}}.$$

1. Merci à Sébastien Gerekens, dont les remarques pertinentes m'ont permis d'optimiser cet énoncé.

En remarquant que

$$\frac{a_{n-1}^{m_1}\cdots a_0^{m_n}}{a_n^{m_1+\cdots+m_n}} = a_n^{d-(m_1+\cdots+m_n)}\,\frac{a_{n-1}^{m_1}\cdots a_0^{m_n}}{a_n^d},$$

on voit qu'il suffit de poser

$$h = \sum_{m_1+\cdots+m_n\leqslant d} c_{m_1,\ldots,m_n}(-1)^{m_1+2m_2+\cdots+nm_n}T_n^{d-(m_1+\cdots+m_n)}T_{n-1}^{m_1}\cdots T_0^{m_n}$$

pour conclure. Cela achève la démonstration. □

1.10. Exemple. Soit $P = a_2X^2 + a_1X + a_0 \in K[X]$, et soient α_1, α_2 les deux racines de P (dans une extension convenable). Par le théorème précédent, on devrait avoir

$$\alpha_1^2 + \alpha_2^2 \in K,$$

et en fait $a_2^2(\alpha_1^2 + \alpha_2^2)$ devrait être un polynôme en a_0, a_1, a_2 à coefficients dans le sous-anneau engendré par 1_K. Vérifions-le. On a

$$\alpha_1^2 + \alpha_2^2 = (\alpha_1 + \alpha_2)^2 - 2\alpha_1\alpha_2 = \sigma_1(\alpha_1, \alpha_2)^2 - 2\sigma_2(\alpha_1, \alpha_2).$$

Par la proposition 1.8, on obtient

$$\alpha_1^2 + \alpha_2^2 = \frac{a_1^2}{a_2^2} - 2\frac{a_0}{a_2} = \frac{a_1^2 - 2a_0a_2}{a_2^2}.$$

On introduit maintenant le discriminant d'un polynôme.

1.11. Proposition. *Soit $P = a_nX^n + \cdots + a_1X + a_0 \in K[X]$ un polynôme de degré $n \geqslant 1$, et soient $\alpha_1, \ldots, \alpha_n$ les racines de P dans une extension convenable L/K. Alors, l'élément*

$$\operatorname{disc}(P) = a_n^{2(n-1)} \prod_{1\leqslant i<j\leqslant n} (\alpha_j - \alpha_i)^2$$

est un élément de K, indépendant du choix de L. C'est même une expression polynomiale en les coefficients de P à coefficients dans le sous-anneau A_0 engendré par 1_K. De plus, $\operatorname{disc}(P) \neq 0$ si, et seulement si, P n'a que des racines simples dans L.

Démonstration. Posons

$$\Delta = \prod_{1\leqslant i<j\leqslant n} (X_j - X_i)^2 \in A_0[X_1, \ldots, X_n].$$

Alors, Δ est symétrique.

En effet, on considère la matrice de Vandermonde

$$M = \left(X_j^{i-1}\right)_{1\leqslant i,j\leqslant n} \in \mathrm{M}_n\big(K(X_1,\ldots,X_n)\big).$$

Il est bien connu (cf. exercice IV-1) que l'on a

$$\det(M) = \prod_{1\leqslant i<j\leqslant n} (X_j - X_i).$$

Ainsi, $\Delta = \det(M)^2$. Soit $\tau \in \mathfrak{S}_n$ une permutation. Alors, on a

$$\Delta(X_{\tau(1)},\ldots,X_{\tau(n)}) = \det(M_\tau)^2,$$

où

$$M_\tau = \left(X_{\tau(j)}^{i-1}\right)_{1\leqslant i,j\leqslant n},$$

puisqu'une permutation des variables induit la même permutation sur les colonnes de M. Comme M_τ est obtenu de M par permutation des colonnes, on a

$$\det(M_\tau) = \pm\det(M),$$

et par conséquent

$$\Delta(X_{\tau(1)},\ldots,X_{\tau(n)}) = \det(M_\tau)^2 = \det(M)^2 = \Delta(X_1,\ldots,X_n).$$

Ainsi, Δ est un polynôme symétrique de degré partiel $2(n-1)$. Par le théorème précédent, l'élément

$$\mathrm{disc}(P) = a_n^{2(n-1)}\Delta(\alpha_1,\ldots,\alpha_n) = a_n^{2(n-1)} \prod_{1\leqslant i<j\leqslant n} (\alpha_j - \alpha_i)^2$$

est un polynôme en les coefficients de P à coefficients dans A_0.

Par définition, $\mathrm{disc}(P) \neq 0$ si, et seulement si, $\alpha_1,\ldots,\alpha_n$ sont deux à deux distincts. □

1.12. Définition. L'élément $\mathrm{disc}(P) \in K$ est appelé le *discriminant* de P.

1.13. Exemples

(1) Si $P = a_2X^2 + a_1X + a_0 \in K[X]$, on a

$$\mathrm{disc}(P) = a_2^2(\alpha_2 - \alpha_1)^2 = a_2^2\big((\alpha_1+\alpha_2)^2 - 4\alpha_1\alpha_2\big).$$

Comme

$$\alpha_1 + \alpha_2 = \sigma_1(\alpha_1,\alpha_2) = -\frac{a_1}{a_2},$$

et

$$\alpha_1\alpha_2 = \sigma_2(\alpha_1,\alpha_2) = \frac{a_0}{a_2},$$

on obtient

$$\operatorname{disc}(P) = a_2^2\Big(\frac{a_1^2}{a_2^2} - 4\frac{a_0}{a_2}\Big) = a_1^2 - 4a_2a_0,$$

qui est la définition usuelle du discriminant d'un polynôme de degré 2.

(2) Soit $P = a_3X^3 + a_2X^2 + a_1X + a_0 \in K[X]$. En appliquant l'algorithme donné dans la démonstration du théorème fondamental sur les polynômes symétriques, on montre que l'on a

$$\Delta = -4\sigma_1^3\sigma_3 + \sigma_1^2\sigma_2^2 - 4\sigma_2^3 + 18\sigma_1\sigma_2\sigma_3 - 27\sigma_3^2.$$

En particulier, si $P = X^3 + pX + q$, on a

$$\sigma_1(\alpha_1,\alpha_2,\alpha_3) = 0, \quad \sigma_2(\alpha_1,\alpha_2,\alpha_3) = p, \quad \sigma_3(\alpha_1,\alpha_2,\alpha_3) = -q,$$

et ainsi

$$\operatorname{disc}(P) = -4p^3 - 27q^2,$$

qui est aussi la quantité bien connue permettant de décider si P a des racines multiples ou non.

Comme dernière application, nous démontrons maintenant le théorème de d'Alembert.

1.14. Théorème (d'Alembert). *Tout polynôme non constant de $\mathbb{C}[X]$ admet une racine dans $\mathbb{C}$.*

Démonstration. Commençons par le cas des polynômes $P \in \mathbb{R}[X]$. Sans perte de généralité, on peut supposer P unitaire. On montre alors le résultat par récurrence sur $n = v_2\big(\deg(P)\big)$. Pour tout $n \geqslant 0$, soit (H_n) la propriété suivante :

(H_n) tout polynôme $P \in \mathbb{R}[X]$ unitaire non constant à coefficients réels tel que $v_2\big(\deg(P)\big) \leqslant n$ admet une racine dans $\mathbb{C}$.

Supposons tout d'abord que $n = 0$. Autrement dit, P est unitaire de degré impair. Dans ce cas, on a

$$\lim_{+\infty} P = +\infty \quad \text{et} \quad \lim_{-\infty} P = -\infty,$$

la seconde égalité utilisant le fait que le degré de P est impair. En particulier, il existe a, $b \in \mathbb{R}$ tels que $P(a) < 0$ et $P(b) > 0$. D'après le théorème des valeurs intermédiaires, P admet une racine réelle. En particulier, (H_0) est vraie. Supposons que (H_n) soit vraie pour un entier $n \geqslant 0$, et soit $P \in \mathbb{R}[X]$ un polynôme unitaire non constant tel que $v_2\big(\deg(P)\big) \leqslant n+1$. Écrivons alors $\deg(P) = d = 2^k s$, où s est impair et $k \in [\![0, n+1]\!]$. Si $k = 0$, on applique (H_0), et l'on a fini. On peut donc supposer que $k \geqslant 1$.

Soit $L/\mathbb{C}$ une extension dans laquelle P est scindé. On a donc

$$P = (X - \alpha_1) \cdots (X - \alpha_d), \quad \alpha_i \in L.$$

Soit $c \in \mathbb{R}$ un réel fixé. Posons

$$Q = \prod_{i<j} \big(X - (\alpha_i + \alpha_j - c\alpha_i\alpha_j)\big) \in L[X].$$

Les coefficients de Q sont des polynômes réels symétriques en $\alpha_1, \ldots, \alpha_d$. Par le théorème 1.9, $Q \in \mathbb{R}[X]$. De plus, on a

$$\deg(Q) = \frac{d(d-1)}{2} = 2^{k-1}s(2^k s - 1).$$

Comme $k \geqslant 1$, on en déduit que

$$v_2\big(\deg(Q)\big) = k - 1 \leqslant n.$$

Par hypothèse de récurrence, Q possède une racine dans $\mathbb{C}$. Il existe donc des entiers $i_c, j_c \in [\![1, d]\!]$, avec $i_c < j_c$, tels que $\alpha_{i_c} + \alpha_{j_c} - c\alpha_{i_c}\alpha_{j_c} \in \mathbb{C}$. Comme $\mathbb{R}$ est infini et que l'ensemble $\{(i,j) \mid 1 \leqslant i < j \leqslant d\}$ est fini, il existe deux réels c_1, c_2 distincts tels que $(i_{c_1}, j_{c_1}) = (i_{c_2}, j_{c_2})$.

Autrement dit, il existe i, $j \in [\![1, d]\!]$, avec $i < j$, et deux réels c_1, c_2 distincts tels que

$$\alpha_i + \alpha_j - c_1\alpha_i\alpha_j, \quad \alpha_i + \alpha_j - c_2\alpha_i\alpha_j \in \mathbb{C}.$$

Par différence, en utilisant le fait que $c_1 - c_2 \in \mathbb{R}^\times$, on obtient $\alpha_i\alpha_j \in \mathbb{C}$. On en déduit ensuite que $\alpha_i + \alpha_j \in \mathbb{C}$. Mais alors, le polynôme $(X - \alpha_i)(X - \alpha_j)$ est un polynôme de degré 2 à coefficients dans $\mathbb{C}$. Or, un tel polynôme a ses deux racines [2] dans $\mathbb{C}$, soit α_i, $\alpha_j \in \mathbb{C}$. Ainsi, P a une racine dans $\mathbb{C}$, et la propriété (H_{n+1}) est vraie. Cela achève la récurrence, et règle le cas des polynômes à coefficients réels.

Revenons au cas général. Si $P = a_d X^d + \cdots + a_1 X + a_0 \in \mathbb{C}[X]$ est non constant, on pose

$$\overline{P} = \overline{a}_d X^d + \cdots + \overline{a}_1 X + \overline{a}_0.$$

Le lecteur vérifiera alors que le polynôme $P\overline{P}$ est à coefficients réels. D'après ce qui précède, $P\overline{P}$ admet une alors une racine $\alpha \in \mathbb{C}$. On a donc $P(\alpha) = 0$ ou $\overline{P}(\alpha) = 0$. Dans le premier cas, on a fini. Dans le second cas, on a donc

$$\overline{a}_d \alpha^d + \cdots + \overline{a}_1 \alpha + \overline{a}_0 = 0.$$

En appliquant la conjugaison complexe à cette égalité, on obtient $P(\overline{\alpha}) = 0$. Cela achève la démonstration. □

2. Il suffit pour cela de montrer que tout nombre complexe a une racine carrée, ce qui s'obtient aisément à partir la forme exponentielle, ou par des calculs purement algébriques.

2. Corps de rupture, corps des racines d'un polynôme

Nous venons de voir qu'étant donné un polynôme $f \in K[X]$ non constant, il existe au moins une extension L/K dans laquelle f a une racine/se scinde. Bien entendu, une telle extension L/K n'est pas unique, puisque toute extension de L convient également. Nous allons maintenant démontrer l'existence d'une extension « minimale » pour cette propriété, et unique à isomorphisme de K-algèbres près.

Soit $f \in K[X]$ un polynôme non constant, et soit L/K une extension dans laquelle f possède une racine $\alpha \in L$. On a donc $K(\alpha) \subset L$, puisque L est une extension de K contenant α.

On remarque alors que $K(\alpha)/K$ est aussi une extension dans laquelle f possède une racine. Cela motive la définition suivante.

2.1. Définition. Soit K un corps, et soit $f \in K[X]$ non constant. Un *corps de rupture* de f est une extension L/K vérifiant les deux conditions suivantes :

(1) le polynôme f possède une racine $\alpha \in L$;

(2) on a $L = K(\alpha)$.

Un tel corps de rupture existe (on choisit une extension de K assez grande dans laquelle f possède une racine, et l'on prend la sous-extension engendrée par cette racine). On va voir que l'on a unicité d'un tel corps de rupture pour un polynôme irréductible unitaire.

Pour cela, nous allons utiliser un résultat central en théorie des corps, et qui nous servira constamment.

2.2. Théorème (de prolongement des morphismes d'anneaux) *Soit L/K une extension, et soit $\alpha \in L$ un élément algébrique sur K. Soit M un corps quelconque, et soit $\iota : K \longrightarrow M$ un morphisme d'anneaux. Alors, l'ensemble des morphismes d'anneaux $\sigma : K(\alpha) \longrightarrow M$ prolongeant ι est en bijection avec l'ensemble des racines*[3] *de $\iota(\mu_{\alpha,K})$ dans M.*

Plus précisément :

(1) *pour tout morphisme d'anneaux $\sigma : K(\alpha) \longrightarrow M$ prolongeant ι, l'élément $\sigma(\alpha)$ est une racine de $\iota(\mu_{\alpha,K})$ dans M ;*

3. S'il en existe !

(2) *pour toute racine $\beta \in M$ de $\iota(\mu_{\alpha,K})$, il existe un unique morphisme d'anneaux $\sigma_\beta : K(\alpha) \longrightarrow M$ tel que*

$$\sigma_\beta(\alpha) = \beta \quad \text{et} \quad \sigma_{\beta|_K} = \iota.$$

Il est défini par

$$\sigma_\beta\big(P(\alpha)\big) = \iota(P)(\beta), \quad \text{pour tout } P \in K[X];$$

(3) *les deux applications*

$$\begin{array}{ccc} \sigma & \longmapsto & \sigma(\alpha) \\ \sigma_\beta & \longmapsto & \beta \end{array}$$

sont bijectives, et inverses l'une de l'autre.

Démonstration. Soit $\sigma : K(\alpha) \longrightarrow M$ un morphisme d'anneaux prolongeant le morphisme $\iota : K \longrightarrow M$. Comme on l'a déjà remarqué en début de chapitre, pour tout $P \in K[X]$, on a

$$\sigma\big(P(\alpha)\big) = \sigma(P)\big(\sigma(\alpha)\big), \quad \text{pour tout } P \in K[X],$$

où $\sigma(P)$ est l'image de P par le morphisme $K[X] \longrightarrow M[X]$ induit par σ. Comme $P \in K[X]$ et que σ prolonge ι, cette égalité se récrit

$$\sigma\big(P(\alpha)\big) = \iota(P)\big(\sigma(\alpha)\big), \quad \text{pour tout } P \in K[X]. \qquad (*)$$

En appliquant cela à $P = \mu_{\alpha,K}$, on obtient que $\sigma(\alpha) \in M$ est une racine de $\iota(\mu_{\alpha,K}) \in M[X]$, d'où (1).

Montrons (2). Soit $\beta \in M$ une racine de $\iota(\mu_{\alpha,K})$. Supposons qu'il existe un morphisme d'anneaux $\sigma_\beta : K(\alpha) \longrightarrow M$ prolongeant ι tel que $\sigma_\beta(\alpha) = \beta$. L'égalité $(*)$ montre alors que

$$\sigma_\beta\big(P(\alpha)\big) = \iota(P)(\beta), \quad \text{pour tout } P \in K[X].$$

Comme α est algébrique sur K par hypothèse, tout élément de $K(\alpha)$ est de la forme $P(\alpha)$, avec $P \in K[X]$, d'après le théorème XXV-3.8. Ainsi, si σ_β existe, il est unique, et défini par la formule précédente. Il reste à démontrer qu'un tel morphisme existe bel et bien.

Puisque $\beta \in M$ est une racine de $\iota(\mu_{\alpha,K})$, le morphisme

$$\begin{array}{ccc} K[X] & \longrightarrow & M \\ P & \longmapsto & \iota(P)(\beta) \end{array}$$

se factorise en un morphisme d'anneaux

$$\begin{array}{rccc} \varphi : & K[X]/(\mu_{\alpha,K}) & \longrightarrow & M \\ & \overline{P} & \longmapsto & \iota(P)(\beta). \end{array}$$

D'après la remarque XXV-3.9, on a un isomorphisme de K-algèbres

$$K[X]/(\mu_{\alpha,K}) \xrightarrow{\sim} K(\alpha),$$

qui envoie $\overline{X}$ sur α. En composant φ et l'inverse de cet isomorphisme, on obtient un morphisme $\sigma_\beta : K(\alpha) \longrightarrow M$ vérifiant les propriétés voulues.

Le point (3) est alors une conséquence immédiate des deux premiers points. Cela achève la démonstration. □

On peut maintenant démontrer l'unicité d'un corps de rupture, lorsque le polynôme $f \in K[X]$ est irréductible.

2.3. Proposition. *Soient K, K' deux corps, soit $\iota : K \xrightarrow{\sim} K'$ un isomorphisme d'anneaux et soit $f \in K[X]$ irréductible unitaire. Soit $K(\alpha)/K$ un corps de rupture de f et soit $K'(\alpha')/K'$ un corps de rupture de $\iota(f)$. Alors, il existe un isomorphisme $K(\alpha) \xrightarrow{\sim} K'(\alpha')$ qui prolonge ι et qui envoie α sur α'.*

En particulier, deux corps de rupture de f sont K-isomorphes.

Démonstration. En appliquant le théorème 2.2 à $M = K'(\alpha')$, on obtient l'existence d'un morphisme d'anneaux $\sigma : K(\alpha) \longrightarrow K'(\alpha')$ prolongeant ι tel que

$$\sigma\big(P(\alpha)\big) = \iota(P)(\alpha'), \qquad \text{pour tout } P \in K[X].$$

Ce morphisme est injectif, et envoie α sur α'. De plus, son image est $K'[\alpha']$, puisque ι est un isomorphisme, d'image K'. Enfin, comme α' est algébrique sur K', on a $K'[\alpha'] = K'(\alpha')$ par le théorème XXV-3.8, d'où le premier point.

Pour obtenir le dernier point, il suffit de prendre $K = K'$ et $\iota = \mathrm{Id}_K$. □

2.4. Exemples

(1) Si $f = X^3 - 2 \in \mathbb{Q}[X]$, un corps de rupture est par exemple $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$. Remarquons que, dans ce cas, f n'a pas toutes ses racines dans un même corps de rupture, puisque f admet des racines complexes non réelles.

(2) Le résultat d'unicité tombe à l'eau si f n'est pas irréductible. Il suffit de prendre par exemple $f = (X^2 + 1)X \in \mathbb{Q}[X]$ pour s'en convaincre. Alors, $\mathbb{Q}/\mathbb{Q}$ est un corps de rupture de f, ainsi que $\mathbb{Q}(i)/\mathbb{Q}$, mais ces extensions ne sont pas isomorphes, car elles n'ont pas le même degré.

Passant maintenant au cas d'une extension de corps L/K dans laquelle un polynôme $f \in K[X]$ non constant donné se scinde. Si $\alpha_1, \ldots, \alpha_n \in L$ sont les racines de f, on a donc $K(\alpha_1, \ldots, \alpha_n) \subset L$, puisque L est une extension de K contenant $\alpha_1, \ldots, \alpha_n$.

On remarque alors que $K(\alpha_1, \ldots, \alpha_n)/K$ est aussi une extension dans laquelle f se décompose en produit de facteurs de degré 1. Cela motive la définition suivante.

2.5. Définition. Soit K un corps, et soit $f \in K[X]$ non constant. Un *corps des racines* de f (ou *corps de décomposition*) est une extension L/K vérifiant les deux conditions suivantes :

(1) le polynôme f se décompose en facteurs de degré 1 dans $L[X]$;

(2) $L = K(\alpha_1, \ldots, \alpha_n)$, où $\alpha_1, \ldots, \alpha_n \in L$ sont les racines de f dans L.

2.6. Remarque. Comme $\alpha_1, \ldots, \alpha_n \in L$ sont algébriques sur K, un corps des racines L/K de f est algébrique de degré fini (cf. corollaire XXV-3.11).

Un tel corps des racines existe (on prend une extension de K assez grande dans laquelle f se scinde et l'on prend la sous-extension engendrée par les racines de f).

Nous allons maintenant démontrer l'unicité d'un corps des racines, à isomorphisme de K-algèbres près.

2.7. Proposition. *Soient K, K' deux corps, soit $\iota : K \xrightarrow{\sim} K'$ un isomorphisme d'anneaux et soit $f \in K[X]$ non constant. Soit L/K un corps de racines de f, et soit L'/K' un corps de racines de $\iota(f)$. Alors, il existe un isomorphisme $\varphi : L \xrightarrow{\sim} L'$ qui prolonge ι.*

En particulier, deux corps de racines de f sont K-isomorphes.

Démonstration. On va démontrer par récurrence la propriété un peu plus générale suivante :

(H_n) le résultat est vrai pour toutes extensions M/K et M'/K', tout isomorphisme $\iota : M \xrightarrow{\sim} M'$, tout $f \in M[X]$ non constant et tous corps des racines L/M et L'/M' de $\iota(f)$ tels que $1 \leqslant [L : M] \leqslant n$.

Démontrons (H_1). Si $n = 1$, alors $L = M$. Dans ce cas, f a toutes ses racines dans M, donc $\iota(f)$ a toutes ses racines dans $\iota(M) = M'$. En effet, si

$$f = c(X - \lambda_1) \cdots (X - \lambda_m), \quad c,\ \lambda_1, \ldots, \lambda_m \in M,\ c \neq 0,$$

alors

$$\iota(f) = \iota(c)\big(X - \iota(\lambda_1)\big) \cdots \big(X - \iota(\lambda_n)\big) \in M'[X].$$

Par définition d'un corps des racines, on a alors

$$L' = M'\big(\iota(\lambda_1), \ldots, \iota(\lambda_m)\big) = M',$$

car chaque $\iota(\lambda_j) \in M'$. Ainsi $L' = \iota(M)$ et l'on prend $\varphi = \iota$.

Supposons (H_n) vraie et montrons (H_{n+1}). Si $[L : M] \leqslant n$, on applique l'hypothèse de récurrence. On peut donc supposer que $[L : M] = n + 1$. Alors $[L : M] \geqslant 2$, donc f admet une racine $\alpha \in L \setminus M$.

Soit $g = \mathrm{Irr}(\alpha, M)$. Alors $\deg(g) \geqslant 2$. Puisque $f(\alpha) = 0$, $g \mid f$. C'est donc un facteur irréductible de f dans $M[X]$.

On constate facilement que $\iota(g)$ est aussi un facteur irréductible de $\iota(f)$ dans $M'[X]$ de même degré. Comme le polynôme $\iota(f)$ se décompose en produit de facteurs de degré 1 dans $L'[X]$, $\iota(g)$ aussi. Soit alors $\alpha' \in L'$ une racine de $\iota(g)$ dans L'. Posons $E = M(\alpha)$ et $E' = M'(\alpha')$. Ce sont des corps de rupture de g et $\iota(g)$ respectivement. Par la proposition 2.3, il existe un isomorphisme d'anneaux $\psi : E \xrightarrow{\sim} E'$ qui prolonge ι et qui envoie α sur α'.

Remarquons maintenant que l'on a $f \in M[X] \subset E[X]$ et $X - \alpha \in E[X]$. Comme on a $f(\alpha) = 0$, on peut donc écrire

$$f = (X - \alpha)P, \quad P \in E[X].$$

De plus, $\iota(g)(\alpha') = 0$. Comme $g \mid f$, on a $\iota(g) \mid \iota(f)$, on a aussi $\iota(f)(\alpha') = 0$, et ainsi $\iota(f) = (X - \alpha')Q$, $Q \in E'[X]$.

Comme $g = \mathrm{Irr}(\alpha, M)$ est de degré $\geqslant 2$, on a

$$[E : M] = [M(\alpha) : M] = \deg(g) \geqslant 2,$$

et la formule de multiplicativité des degrés montre que $[L : E] < [L : M]$. Or, L/E est un corps des racines de P. En effet, si $\alpha, \alpha_2, \ldots, \alpha_m$ sont les racines de f dans L, alors $\alpha_2, \ldots, \alpha_m$ sont les racines de $P \in E[X]$ dans L, et l'on a

$$L = M(\alpha, \alpha_2, \ldots, \alpha_m) = M(\alpha)(\alpha_2, \ldots, \alpha_m) = E(\alpha_2, \ldots, \alpha_m).$$

De même L'/E' est un corps des racines de Q.

Comme ψ prolonge ι, on a l'égalité $\iota(f) = \psi(f)$ dans $E[X]$. Ainsi, on a

$$\iota(f) = \psi(X - \alpha)\psi(P) = \big(X - \psi(\alpha)\big)\psi(P) = (X - \alpha')\psi(P),$$

car $\psi(\alpha) = \alpha'$. Comme on a aussi $\iota(f) = (X - \alpha')Q$, on a $\psi(P) = Q$ par intégrité de $E'[X]$. Par hypothèse de récurrence, il existe un isomorphisme d'anneaux $\varphi : L \xrightarrow{\sim} L'$ qui prolonge ψ, donc qui prolonge aussi ι. Cela achève la récurrence.

On prend alors $M = K, M' = K'$ pour obtenir le résultat. Pour démontrer la dernière partie, il suffit de prendre $K = K'$ et $\iota = \mathrm{Id}_K$. □

En guise d'application, nous allons retrouver les résultats sur les corps finis obtenus dans le paragraphe XXI-6 d'une autre manière.

2.8. Théorème. *Soit $q \geqslant 2$ une puissance d'un nombre premier p. Alors, il existe un corps fini à q éléments, unique à isomorphisme de $\mathbb{F}_p$-algèbres près. C'est le corps des racines du polynôme $X^q - X \in \mathbb{F}_p[X]$.*

Démonstration. On commence par l'existence. Soit $L/\mathbb{F}_p$ un corps des racines de $f = X^q - X \in \mathbb{F}_p[X]$. Montrons que $X^q - X$ admet q racines distinctes dans L. Si f a une racine multiple α, on a

$$X^q - X = (X - \alpha)^2 g, \quad g \in L[X].$$

En dérivant formellement, on obtient,

$$qX^{q-1} - 1 = -1 = (X - \alpha)h, \quad h \in L[X],$$

d'où une contradiction en prenant les degrés (le corps L est aussi de caractéristique p !).

Ainsi, f admet q racines distinctes. Montrons que l'ensemble K de ces racines forment un corps. D'après le lemme XV-1.18, l'application

$$\begin{aligned} L &\longrightarrow L \\ x &\longmapsto x^q \end{aligned}$$

est un morphisme d'anneaux. Comme K est par définition l'ensemble des éléments de L fixés par ce morphisme d'anneaux, c'est un sous-corps de L par le lemme XXV-4.10.

Montrons maintenant l'unicité. Soit K un corps à q éléments. Comme K est de caractéristique p, on a $\mathbb{F}_p \subset K$. On peut donc considérer l'ensemble des racines de $f = X^q - X \in \mathbb{F}_p[X]$. Comme $K^\times$ est un groupe multiplicatif d'ordre $q - 1$, par le théorème de Lagrange, on a

$$x^{q-1} = 1, \quad \text{pour tout } x \in K^\times.$$

On a donc $x^q = x$ pour tout $x \in K$. Ainsi, les q éléments de K sont racines de $f \in \mathbb{F}_p[X]$. Comme $X^q - X$ possède au plus q racines dans K, f possède exactement q racines dans K, et ces racines sont les éléments de K. Remarquons aussi que puisque K contient $\mathbb{F}_p$ et les x_i, K contient $\mathbb{F}_p(x_1, \ldots, x_q)$. L'autre inclusion étant évidente, on a $K = \mathbb{F}_p(x_1, \ldots, x_q)$. Ainsi, K est un corps des racines de $X^q - X$. Par conséquent, il est unique à isomorphisme de $\mathbb{F}_p$-algèbres près. □

On peut alors retrouver le fait qu'il existe des polynômes irréductibles de tout degré sur $\mathbb{F}_p$ (cf. proposition XXI-6.7). Comme ça ne coûte pas plus cher, nous allons montrer que le résultat vaut sur $\mathbb{F}_q$. On commence par un lemme.

2.9. Lemme. *Soit K un corps fini, et soit L/K une extension de degré fini. Alors, il existe $\alpha \in L$ tel que $L = K(\alpha)$.*

Démonstration. Les hypothèses impliquent que L est aussi un corps fini. En particulier, le groupe $L^\times$ est cyclique par le théorème XII-2.8. Soit $\alpha \in L^\times$ un générateur de $L^\times$, et soit $n \geqslant 1$ son ordre. On a alors

$$L = \{0, 1, \alpha, \ldots, \alpha^{n-1}\} \subset K(\alpha).$$

Comme L contient α et K, on en déduit que $K(\alpha) \subset L$, d'où $L = K(\alpha)$.□

2.10. Corollaire. *Soit $q \geqslant 2$ une puissance d'un nombre premier p. Alors, pour tout entier $d \geqslant 1$, il existe au moins un polynôme irréductible unitaire de $\mathbb{F}_q[X]$ de degré d.*

Démonstration. Soit L le corps des racines de $X^{q^d} - X$ sur $\mathbb{F}_p$. C'est un corps fini à q^d éléments par le théorème 2.8. Or, $X^q - X \mid X^{q^d} - X$ dans $\mathbb{F}_p[X]$, puisque par exemple dans l'anneau $\mathbb{F}_p[X]/(X^q - X)$, on a

$$\overline{X^{q^d}} = \overline{X}^{q^d} = \overline{X^q}^{q^{d-1}} = \overline{X}^{q^{d-1}} = \cdots = \overline{X}.$$

En particulier, $X^q - X$ est scindé dans L, et L contient un corps des racines K de $X^q - X$ sur $\mathbb{F}_p$, qui est un corps fini à q éléments. Mais alors, L/K est une extension de degré d. Par le lemme 2.9, on a $L = K(\alpha)$. Le polynôme $\mu_{\alpha,K}$ est alors un polynôme unitaire de degré d. Comme $K \simeq \mathbb{F}_q$ par le théorème 2.8, on a fini. □

On vient de voir dans la démonstration précédente que $\mathbb{F}_{q^d}$ contient le corps $\mathbb{F}_q$. Le lemme 2.9 et le théorème 2.2 permettent alors de calculer le groupe des automorphismes de $\mathbb{F}_{q^d}/\mathbb{F}_q$.

2.11. Proposition. *Soit $q \geqslant 2$ une puissance d'un nombre premier p. Alors, pour tout $d \geqslant 1$, le groupe $\mathrm{Gal}(\mathbb{F}_{q^d}/\mathbb{F}_q)$ est cyclique d'ordre $d \geqslant 1$, engendré par l'automorphisme de Frobenius*

$$\begin{aligned} \sigma : \mathbb{F}_{q^d} &\longrightarrow \mathbb{F}_{q^d} \\ x &\longmapsto x^q. \end{aligned}$$

Démonstration. Commençons par montrer que $\sigma \in \mathrm{Gal}(\mathbb{F}_{q^d}/\mathbb{F}_q)$. Comme q est une puissance d'un nombre premier p, qui est la caractéristique de $\mathbb{F}_q$ et de $\mathbb{F}_{q^d}$, le lemme XV-1.18 montre que σ est un morphisme d'anneaux. C'est donc un isomorphisme par le lemme XXI-6.9. Enfin, on a $\sigma(x) = x$ pour tout $x \in \mathbb{F}_q$, par le lemme XXI-6.3. Cela implique facilement que σ est $\mathbb{F}_q$-linéaire.

Soit $m \geqslant 1$ l'ordre de σ. On a alors

$$\sigma^m = \mathrm{Id}_{\mathbb{F}_{q^d}} \iff x^{q^m} = x, \text{ pour tout } x \in \mathbb{F}_{q^d}.$$

En particulier, le polynôme $X^{q^m} - X \in \mathbb{F}_q[X]$ de degré q^m possède au moins q^d racines, et ainsi $m \geqslant d$. Par ailleurs, on a aussi $\sigma^d = \mathrm{Id}_{\mathbb{F}_q}$ par le lemme XXI-6.3. Ainsi, σ est d'ordre d. Il reste à démontrer que σ engendre le groupe $\mathrm{Gal}(\mathbb{F}_{q^d}/\mathbb{F}_q)$. D'après le lemme 2.9, on a $\mathbb{F}_{q^d} = \mathbb{F}_q(\alpha)$, pour un certain $\alpha \in \mathbb{F}_{q^d}$. Le théorème 2.2 appliqué à $\iota = \mathrm{Id}_{\mathbb{F}_q}$ montre en particulier que $\mathrm{Gal}(\mathbb{F}_{q^d}/\mathbb{F}_q)$ possède au plus d éléments. Comme ce groupe contient $\langle\sigma\rangle$, qui possède d éléments, on obtient le résultat souhaité. □

3. Clôture algébrique d'un corps

Nous allons maintenant nous intéresser à la construction d'une extension de corps L/K « contenant tous les éléments algébriques sur K et seulement eux ». Bien sûr, cette formulation est pour l'instant très imprécise. De manière plus rigoureuse, l'extension L/K doit être algébrique, et tout polynôme de $K[X]$ non constant doit avoir une racine dans L.

Si elle existe, une telle extension semble vérifier les propriétés souhaitées. Néanmoins, rien n'empêche a priori que l'on ait créé de nouveaux éléments algébriques sur K. Par exemple, si E/L est algébrique, alors E/K est aussi algébrique par la proposition XXV-3.19. Si E/L est non triviale, les éléments de $E \setminus L$ sont alors des éléments algébriques sur K non contenus dans L. Nous démontrerons plus loin que ce phénomène ne se produit pas, i.e. qu'une telle extension L/K n'a pas d'extension algébrique non triviale.

On commence par quelques définitions.

3.1. Définition. On dit qu'un corps L est *algébriquement clos* si tout polynôme $f \in L[X]$ non constant possède une racine dans L. Une récurrence facile montre que cela revient à dire que tout polynôme $f \in L[X]$ est scindé dans L.

On dit que L est une *clôture algébrique* de K si L/K est une extension algébrique et L est algébriquement clos.

3.2. Exemple. Le corps $\mathbb{C}$ est algébriquement clos d'après le théorème de d'Alembert, mais $\mathbb{C}$ n'est pas une clôture algébrique de $\mathbb{Q}$, car il contient au moins un élément transcendant sur $\mathbb{Q}$. En revanche, $\mathbb{C}$ est une clôture algébrique de $\mathbb{R}$.

3.3. Remarque. Si L est un corps algébriquement clos, alors l'extension triviale L/L est la seule extension algébrique de L.

En effet, si L'/L est algébrique, et si $\alpha \in L'$, le polynôme minimal de α sur L admet une racine dans L. Comme il est irréductible, il est nécessairement de degré 1. Autrement dit, $\alpha \in L$ et $L' = L$.

Nous avons alors le théorème suivant, qui démontre[4] que l'extension que nous recherchions en début de paragraphe est exactement une clôture algébrique de K.

3.4. Théorème. *Une extension L/K est une clôture algébrique de K si, et seulement si, L/K et algébrique et tout polynôme $P \in K[X]$ possède une racine dans L.*

Démonstration. Le sens direct découle des définitions. Supposons maintenant que L/K vérifie les conditions de l'énoncé. Il reste à démontrer que tout polynôme de $L[X]$ non constant a une racine dans L. Soit $f \in L[X]$ un tel polynôme, et soit $E = L(\alpha)$ un corps de rupture de f. Comme les extensions E/L et L/K sont algébriques, E/K est algébrique par la proposition XXV-3.19. En particulier, α est algébrique sur K. Soit $P = \mu_{\alpha,K}$. Par hypothèse, P possède une racine $\beta \in L$. On a donc aussi $P = \mu_{\beta,K}$.

Supposons tout d'abord que K soit un corps fini à q éléments.

Soit $M = K(\beta)$. C'est un corps fini à q^d éléments, où d est le degré de P. Nous allons démontrer l'égalité

$$P = \prod_{k=0}^{d-1} \left(X - \beta^{q^k}\right).$$

Comme $P(\alpha) = 0$, cela entraînera que α est une puissance de β, et donc que $\alpha \in L$.

Considérons l'automorphisme de Frobenius $\sigma : M \longrightarrow M$ donné par la proposition 2.11. On sait que σ est un élément d'ordre d de $\mathrm{Gal}(M/K)$. Ainsi, $\mathrm{Id}_M, \sigma, \ldots, \sigma^{d-1}$ sont deux à deux distincts. Comme $M = K(\beta)$, le théorème 2.2 montre que les éléments

$$\sigma^k(\beta) = \beta^{q^k}, \quad k \in [\![0, d-1]\!]$$

sont des racines de P deux à deux distinctes, ce qui donne immédiatement l'égalité voulue. Cela achève la démonstration du théorème lorsque K est fini.

4. La démonstration du théorème suivant est tirée de l'excellent livre de P. Colmez [20].

On suppose maintenant que K est infini. Soit $L'/K(\alpha)$ une extension dans laquelle $P \in K[X] \subset K(\alpha)[X]$ est scindé.

On peut donc écrire

$$P = \prod_{k=1}^{d} (X - \alpha_i), \quad \alpha_i \in L',$$

avec $\alpha_1 = \alpha$.

Pour $\mathbf{t} = (t_1, \ldots, t_d) \in K^d$, on pose

$$Q_{\mathbf{t}} = \prod_{\rho \in \mathfrak{S}_d} \big(X - (t_1\alpha_{\rho(1)} + \cdots + t_d\alpha_{\rho(d)})\big).$$

Ses coefficients sont symétriques en $\alpha_1, \ldots, \alpha_d$, et l'on a donc $Q_{\mathbf{t}} \in K[X]$ d'après le théorème 1.9. Par hypothèse sur l'extension L/K, il existe donc une permutation $\rho \in \mathfrak{S}_d$ telle que $t_1\alpha_{\rho(1)} + \cdots + t_d\alpha_{\rho(d)} \in L$, soit encore

$$t_{\rho^{-1}(1)}\alpha_1 + \cdots + t_{\rho^{-1}(d)}\alpha_d \in L.$$

Pour tout $\tau \in \mathfrak{S}_d$, on considère l'automorphisme de K-espaces vectoriels

$$\begin{array}{rccc} u_\tau : & K^d & \longrightarrow & K^d \\ & (t_1, \ldots, t_d) & \longmapsto & (t_{\tau^{-1}(1)}, \ldots, t_{\tau^{-1}(d)}). \end{array}$$

Posons maintenant

$$W = \{\mathbf{t} \in K^d \mid t_1\alpha_1 + \cdots + t_d\alpha_d \in L\}.$$

C'est un sous-espace vectoriel de K^d. Ce qui précède montre alors que

$$K^d = \bigcup_{\tau \in \mathfrak{S}_d} u_\tau(W).$$

Or, on a le fait suivant.

Fait. Soit K un corps infini. Alors, un K-espace vectoriel E n'est pas réunion finie de sous-espaces stricts.

Supposons le fait précédent établi. On en déduit alors qu'il existe une permutation $\tau \in \mathfrak{S}_d$ telle que $u_\tau(W) = K^d$. Comme u_τ est un automorphisme de K^d, on obtient alors $u_\tau(W) = u_\tau(K^d)$, puis $W = K^d$. En particulier, on a $\alpha = \alpha_1 \in L$.

Il reste à établir le fait sus-mentionné. Soit K un corps quelconque, et soit E un K-espace vectoriel. On suppose que E est la réunion des sous-espaces stricts $E_1, \ldots, E_r$ $(r \geqslant 1)$. Quitte à éliminer les redondances, on peut supposer qu'aucun d'entre eux n'est inclus dans la réunion des autres. Choisissons alors un vecteur $u \in E_1$ n'appartenant pas à $E_2 \cup \cdots \cup E_r$. En particulier, u est non nul. De plus, comme E_1 est un sous-espace strict de E, il existe un vecteur $v \in E \setminus E_1$. L'ensemble $v + Ku$ est alors disjoint de E_1. Supposons maintenant qu'il existe $\lambda_1, \lambda_2 \in K$ tels que $v+\lambda_1 u$, $v+\lambda_2 u \in E_k$, où $k \in [\![2, r]\!]$. Alors, $(\lambda_1 - \lambda_2)u \in E_k$. On a donc $\lambda_1 - \lambda_2 = 0$, car sinon on aurait $u \in E_k$, ce qui est exclu par choix de u.

Autrement dit, $v + Ku$ contient au plus un vecteur de E_k, et cela pour tout entier $k \in [\![2, r]\!]$. Comme E est la réunion de $E_1, \ldots, E_r$, on en déduit que $v + Ku$ contient au plus $r - 1$ éléments. Mais alors, cela implique que le corps K est fini (car u est non nul). Cela achève la démonstration. □

Il est temps de s'intéresser à l'existence et à l'unicité d'une clôture algébrique. À cette fin, commençons par établir une proposition qui montrent que les corps algébriquement clos ont de bonnes propriétés vis-à-vis du prolongement de morphismes.

3.5. Proposition. *Soit L/K une extension algébrique (non nécessairement de degré fini) et soit L' un corps algébriquement clos.*

Alors, tout morphisme d'anneaux $\iota : K \longrightarrow L'$ se prolonge en un morphisme d'anneaux $\varphi : L \longrightarrow L'$.

Démonstration. La démonstration repose sur le lemme de Zorn (et donc l'axiome du choix). Soit $\mathcal{F}$ l'ensemble

$$\mathcal{F} = \{(F, \tau) \mid K \subset F \subset L,\ \tau : F \hookrightarrow L',\ \tau_{|K} = \iota\}.$$

Remarquons que $\mathcal{F}$ est non vide puisqu'il contient (K, ι). On vérifie aisément que la relation $\preceq$ définie sur $\mathcal{F}$ par

$$(F, \tau) \preceq (F', \tau') \quad \Longleftrightarrow \quad F \subset F' \text{ et } \tau'_{|F} = \tau,$$

est une relation d'ordre (partiel) sur $\mathcal{F}$. Montrons que $\mathcal{F}$ est inductif.

Soit $(I, \leqslant)$ un ensemble non vide totalement ordonné, et soit $\big((F_i, \tau_i)\big)_{i \in I}$ une chaîne d'éléments de $\mathcal{F}$ indexée par I. On a donc $(F_i, \tau_i) \preceq (F_j, \tau_j)$ pour tout $i \leqslant j$.

Posons $F = \bigcup_{i \in I} F_i$ et soit $\tau : F \longrightarrow L'$ l'application définie par

$$\tau(x) = \tau_i(x),$$

dès que $x \in F_i$. On vérifie que F est un sous-corps de L contenant K (cela provient du fait que les sous-corps F_i sont emboîtés). De plus, $\tau(x)$ ne dépend pas de l'indice $i \in I$ choisi, puisque, si $x \in F_i \cap F_j$, avec par exemple $i \leqslant j$, alors on a $F_i \subset F_j$ et donc

$$\tau_j(x) = (\tau_j)_{|F_i}(x) = \tau_i(x).$$

On vérifie alors que τ est un morphisme d'anneaux prolongeant ι. Ainsi, on a $(F, \tau) \in \mathcal{F}$. Mais par définition, $F_i \subset F$ pour tout $i \in I$ et $\tau_{|F_i} = \tau_i$. Par conséquent, on a

$$(F_i, \tau_i) \preceq (F, \tau), \qquad \text{pour tout } i \in I.$$

Autrement dit, (F, τ) est un majorant de $\big((F_i, \tau_i)\big)_{i\in I}$, et $\mathcal{F}$ est inductif. Comme il est non vide, il possède un élément maximal (F_0, φ). Pour conclure, il reste donc à montrer que $F_0 = L$.

Supposons a contrario que $F_0 \subsetneq L$, et soit $\alpha \in L \setminus F_0$. Puisque L/K est algébrique, α est algébrique sur K, donc sur F_0. Soit $f = \mathrm{Irr}(\alpha, F_0)$. Puisque L' est algébriquement clos, $\varphi(f) \in L'[X]$ possède une racine dans L'. Le théorème de prolongement des morphismes montre alors qu'il existe un morphisme d'anneaux $\psi : F_0(\alpha) \longrightarrow L'$ qui prolonge φ (et donc ι). De plus, on a $F_0 \subsetneq F_0(\alpha)$ puisque $\alpha \notin F_0$. Ainsi, $(F_0, \varphi) \prec (F_0(\alpha), \psi)$, ce qui contredit la maximalité de (F_0, φ). On a donc $F_0 = L$, et cela achève la démonstration. □

On peut maintenant démontrer l'existence et l'unicité d'une clôture algébrique.

3.6. Théorème (Steinitz). *Tout corps K admet une clôture algébrique, unique à isomorphisme de K-algèbres près.*

Démonstration. Commençons par démontrer l'unicité d'une clôture algébrique. Soient L_1 et L_2 deux clôtures algébriques de K. Puisque l'extension L_1/K est algébrique et que le corps L_2 est algébriquement clos, la proposition précédente montre que l'inclusion $K \subset L_2$ se prolonge en un morphisme d'anneaux $\sigma : L_1 \longrightarrow L_2$, qui est alors K-linéaire, et donc un morphisme de K-algèbres. De plus, σ étant injectif, il reste à démontrer sa surjectivité. Remarquons tout d'abord que $\sigma(L_1)$ est algébriquement clos. En effet, soit $g \in \sigma(L_1)[X]$ non constant. Alors, on peut écrire

$$g = \sigma(a_n)X^n + \cdots + \sigma(a_1)X + \sigma(a_0), \quad a_i \in L_1,\ a_n \neq 0,\ n \geqslant 1.$$

Mais alors $g = \sigma(f)$, où $f = a_nX^n + \cdots + a_1X + a_0 \in L[X]$. Comme f est aussi non constant, il possède une racine $\alpha \in L_1$, et ainsi $\sigma(\alpha)$ est une racine de g dans $\sigma(L_1)$.

Soit maintenant $\beta \in L_2$. Comme β est algébrique sur K, il existe $f \in K[X]$ non nul tel que $f(\beta) = 0$. Comme $f \in K[X] \subset \sigma(L_1)[X]$ est non constant (puisqu'il a une racine et qu'il est non nul) et que $\sigma(L_1)$ est algébriquement clos, f possède toutes ses racines dans $\sigma(L_1)$. En particulier, $\beta \in \sigma(L_1)$ et l'on a donc $\sigma(L_1) = L_2$, d'où la surjectivité de σ.

Passons maintenant à l'existence d'une clôture algébrique. Pour tout polynôme $f \in K[T] \setminus K$, on considère une indéterminée X_f, et l'on considère l'anneau des polynômes $A = K[(X_f)_{f\in K[T]\setminus K}]$. Autrement dit, c'est la réunion des anneaux de polynômes $K[X_{f_1}, \ldots, X_{f_n}]$, où $n \geqslant 0$ est un entier et $f_1, \ldots, f_n \in K[T] \setminus K$.

On considère $\mathfrak{a}$ l'idéal de A engendré par les éléments

$$f(X_f), f \in K[T] \setminus K.$$

Montrons que $\mathfrak{a} \neq A$. Dans le cas contraire, on aurait $1 \in \mathfrak{a}$. Il existerait alors des polynômes $f_1, \ldots, f_n \in K[T]$ et des éléments $h_1, \ldots, h_n \in A$ tels que

$$1 = h_1 f_1(X_{f_1}) + \cdots + h_n f(X_{f_n}).$$

Remarquons que chaque f_i possède une racine α_i dans un corps des racines de $f_1 \cdots f_n \in K[X]$. En spécialisant X_{f_i} en α_i, on obtient alors la contradiction $1 = 0$.

Ainsi, $\mathfrak{a} \neq A$, et il existe un idéal maximal $\mathfrak{m}$ de A contenant $\mathfrak{a}$ par le lemme de Krull. L'anneau quotient $L' = A/\mathfrak{m}$ est alors un corps. Considérons le morphisme de K-algèbres $\iota : K \longrightarrow L'$. Pour tout $f \in K[X]$ non constant, le polynôme $\iota(f)$ a une racine dans L'. En effet, on a

$$\iota(f)(\overline{X}_f) = \overline{f(X_f)} = \overline{0},$$

car $f(X_f) \in \mathfrak{a} \subset \mathfrak{m}$. Par transport de structure, on obtient une extension de corps L/K dans laquelle tout polynôme non constant de $K[X]$ possède une racine. Soit alors $\overline{K}$ l'ensemble des éléments de L algébriques sur K. C'est une sous-extension algébrique de L/K. Si de plus $P \in K[X]$ est un polynôme non constant, alors P admet une racine $\alpha \in L$. Mais, par construction, α est algébrique sur K et donc $\alpha \in \overline{K}$. L'extension $\overline{K}/K$ est alors une clôture algébrique de K, d'après le théorème 3.4. Cela achève la démonstration. □

3.7. Remarque. On peut se passer de l'utilisation du théorème 3.4 afin de démontrer l'existence d'une clôture algébrique de K.

Comme dans la démonstration du théorème de Steinitz, on commence par construire une extension L_1/K dans laquelle tout polynôme non constant de $K[X]$ possède une racine. Par récurrence, on construit une suite $(L_i)_{i \geqslant 1}$ d'extensions de L vérifiant $L_i \subset L_{i+1}$ pour tout $i \geqslant 1$, et telles que tout polynôme non constant de $L_i[X]$ possède une racine dans L_{i+1}. Posons alors $L = \bigcup_{i \geqslant 1} L_i$, et montrons que L est algébriquement clos. Soit un polynôme $f \in L[X]$ non constant. Alors, puisque la suite $(L_i)_{i \geqslant 1}$ est croissante, il existe un indice $j \geqslant 1$ tel que $f \in L_j[X]$. Mais alors par construction, f possède une racine dans L_{j+1}, donc dans L.

Soit $\overline{K}$ l'ensemble des éléments de L algébriques sur K. C'est une sous-extension algébrique de L/K. Si $f \in \overline{K}[X] \subset L[X]$ est un polynôme non constant, f admet une racine $\alpha \in L$. Mais alors, $\overline{K}(\alpha)/\overline{K}$ est algébrique, et ainsi $\overline{K}(\alpha)/K$ est algébrique par la proposition XXV-3.19. En particulier, α

est algébrique sur K, et ainsi $\alpha \in \overline{K}$. Par conséquent, $\overline{K}$ est algébriquement clos, et $\overline{K}/K$ est une clôture algébrique de K.

Dorénavant, nous pouvons fixer une clôture algébrique de K, et faire tous les calculs dans cette extension lorsque l'on travaille avec des extensions de degré fini ou algébriques.

4. Entiers algébriques

Revenons maintenant sur une question soulevée dans le chapitre XXI, à savoir : soit A un anneau intègre, de corps des fractions K_A. Un polynôme unitaire $f \in A[X]$ irréductible est-il encore irréductible dans $K_A[X]$?

Nous avons vu que ce n'était pas toujours le cas (cf. exemple XXI-2.18), mais que cela devient vrai si A est factoriel (cf. remarque XXI-3.11). Notre but est de comprendre à quelles conditions sur A la question précédente admet une réponse positive.

Pour cela, nous avons besoin de la notion d'élément entier sur A.

Avant tout chose, remarquons que si L/K_A est une extension de corps, un élément $\alpha \in L$ algébrique sur K est racine d'un polynôme non constant à coefficients dans A.

En effet, soit $P \in K_A[X]$ est un polynôme non constant tel que $P(\alpha) = 0$. Si $d \in A \setminus \{0\}$ est un dénominateur commun aux coefficients de P, alors on a $dP \in A[X]$ et $(dP)(\alpha) = 0$.

Par ailleurs, par définition, α est aussi racine d'un polynôme unitaire non constant à coefficients dans K_A. On peut légitimement se demander si α est racine d'un polynôme **unitaire** à coefficients dans A.

Il n'en est rien, comme le montre l'exemple suivant.

4.1. Exemple. Soit $A = \mathbb{Z}$. On a donc $K_A = \mathbb{Q}$. Soit $\alpha = \frac{1}{2}$. C'est un élément algébrique sur $\mathbb{Q}$. Supposons que α soit racine d'un polynôme unitaire non constant à coefficients dans $\mathbb{Z}$. Il existe donc un entier $n \geqslant 1$ et $a_0, \ldots, a_{n-1} \in \mathbb{Z}$ tels que

$$\frac{1}{2^n} + \frac{a_{n-1}}{2^{n-1}} + \cdots + \frac{a_1}{2} + a_0 = 0.$$

En multipliant cette égalité par 2^n, on obtient la contradiction que 1 est pair.

Les éléments algébriques qui sont racines d'un polynôme unitaire à coefficients dans A méritent un nom particulier. Comme cela ne coûte pas plus

cher, nous allons nous placer dans un contexte un petit peu plus général. Cela nous permettra au passage de fournir une jolie application de la notion de module sur un anneau, notion que nous avons sous-utilisée dans cet ouvrage.

4.2. Définition. Soit B un anneau, et soit A un sous-anneau de B. On dit que $x \in B$ est *entier* sur A s'il existe un polynôme $P \in A[X]$ unitaire tel que $P(x) = 0$.

En particulier, tout élément de A est donc entier sur A.

4.3. Exemple. Le réel $\phi = \dfrac{1+\sqrt{5}}{2}$ est entier sur $\mathbb{Z}$, puisque l'on a

$$\phi^2 - \phi - 1 = 0.$$

Le lemme suivant donne diverses caractérisations des éléments entiers.

4.4. Lemme. *Soit A un anneau, et soit B contenant A comme sous-anneau. Pour tout $x \in B$, les conditions suivantes sont équivalentes :*

(1) *l'élément x est entier sur A ;*

(2) *le sous-anneau $A[x]$ engendré par A et x est un A-module de type fini ;*

(3) *le sous-anneau $A[x]$ est contenu dans un sous-anneau M de B qui est un A-module de type fini.*

Démonstration

(1) $\Longrightarrow$ (2). Soit $f \in A[X]$ un polynôme unitaire tel que $f(x) = 0$. Tout élément de $A[x]$ est de la forme $P(x)$, avec $P \in A[X]$. Puisque f est unitaire, on peut effectuer la division euclidienne de P par f, et l'on peut donc écrire

$$P = Qf + R,$$

où $Q, R \in A[X]$ et $\deg(R) \leqslant \deg(f) - 1$.

On en déduit alors $P(x) = R(x)$, et donc $P(x)$ est combinaison linéaire des éléments $1, x, \ldots, x^{n-1}$, où $n = \deg(f)$. Par conséquent, $A[x]$ est engendré comme A-module par $1, x, \ldots, x^{n-1}$, et est donc de type fini.

(2) $\Longrightarrow$ (3). Évident.

(3) $\Longrightarrow$ (1). Soient $x_1, \ldots, x_n$ des générateurs du A-module M. On peut supposer qu'ils sont non nuls. Puisque M est un sous-anneau de B, on a alors $xx_i \in M$, pour tout $i \in [\![1, n]\!]$.

On peut donc écrire

$$xx_i = \sum_{j=1}^{n} a_{ij} \cdot x_j, \quad a_{ij} \in A, \quad \text{pour tout } i \in [\![1, n]\!].$$

Soit $C = (a_{ij}) \in \mathrm{M}_n(A)$. Les équations précédentes s'écrivent

$$(x\mathrm{I}_n - C) \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0.$$

En multipliant par la transposée de la comatrice de $x\mathrm{I}_n - C$, on obtient

$$\det(x\mathrm{I}_n - C)\,\mathrm{I}_n \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = 0.$$

On a donc

$$\det(x\mathrm{I}_n - C)x_i = 0, \quad \text{pour tout } i \in [\![1, n]\!].$$

Puisque les x_i engendrent M, on obtient

$$\det(x\mathrm{I}_n - C)z = 0, \quad \text{pour tout } z \in M.$$

En particulier, pour $z = 1$, on obtient

$$\det(x\mathrm{I}_n - C) = 0.$$

Autrement dit, $\chi_C(x) = 0$. Or, $\chi_C \in A[X]$ est unitaire, et donc $x \in B$ est entier sur A. □

Dans la suite, nous allons établir que l'ensemble des éléments de B entiers sur A est un anneau. On commence par un lemme.

4.5. Lemme. *Soit A un anneau, et soit B contenant A comme sous-anneau. Soient $x_1, \ldots, x_n \in B$. Supposons que pour tout entier $i \in [\![1, n]\!]$, l'élément x_i soit entier sur $A[x_1, \ldots, x_{i-1}]$. Alors, l'anneau $A[x_1, \ldots, x_n]$ est un A-module de type fini, et tout élément de $A[x_1, \ldots, x_n]$ est entier sur A.*

Démonstration. Pour $n \geqslant 1$, soit (H_n) la propriété suivante :

(H_n) Pour tout sous-anneau A' de B, et tous $x_1, \ldots, x_n \in B$ tels que x_i soit entier sur $A'[x_1, \ldots, x_{i-1}]$ pour tout $i \in [\![1, n]\!]$, l'anneau $A'[x_1, \ldots, x_n]$ est un A'-module de type fini.

Nous allons démontrer (H_n) pour tout $n \geqslant 1$ par récurrence sur n, ce qui donnera la première partie du lemme en prenant $A' = A$.

Tout d'abord, (H_1) est vrai d'après le lemme précédent. Supposons maintenant (H_n) soit vraie pour un certain entier $n \geqslant 1$, et montrons que (H_{n+1}) est vraie. Soit A' un sous-anneau de B, et soient $x_1, \ldots, x_{n+1} \in B$ tels que pour tout $i \in [\![1, n]\!]$, l'élément x_i soit entier sur $A'[x_1, \ldots, x_{i-1}]$. Par hypothèse, $x_{n+1} \in B$ est entier sur l'anneau $M = A'[x_1, \ldots, x_n]$. D'après le lemme précédent, $M[x_{n+1}]$ est un M-module de type fini. Soient $y_1, \ldots, y_r$ des générateurs de ce M-module. Par hypothèse de récurrence, l'anneau M est un A'-module de type fini, engendré disons par s éléments $z_1, \ldots, z_s$. Il est alors facile de voir que le A'-module $M[x_{n+1}]$ est engendré par la famille $(y_i z_j)_{i,j}$, ce qui achève la récurrence.

Montrons la dernière partie. Si $x \in A[x_1, \ldots, x_n]$, alors $A[x]$ est un sous-anneau de $A[x_1, \ldots, x_n]$. Comme $A[x_1, \ldots, x_n]$ est un A-module de type fini, x est entier sur A d'après le lemme précédent. Cela achève la démonstration. □

4.6. Corollaire. *Soit A un anneau, et soit B un anneau contenant A comme sous-anneau. Alors, l'ensemble des éléments de B entiers sur A est un sous-anneau de B contenant A.*

Démonstration. Soit C l'ensemble des éléments de B entiers sur A. Alors, C contient A. Soient maintenant $x, y \in C$. Puisque y est entier sur A, il est entier sur $A[x]$. D'après le lemme 4.5, tous les éléments de $A[x, y]$ sont entiers sur A. En particulier 0, 1, $x \pm y$ et xy sont entiers sur A, et C est donc un sous-anneau de B. □

4.7. Remarque. Dans les applications qui vont suivre et les exercices, nous nous restreindrons souvent à la situation suivante : on considère un anneau intègre A, et l'on travaille dans une extension L/K_A.

On peut alors démontrer que l'ensemble B des éléments de L entiers sur A est un sous-anneau de L contenant A, sans utiliser la notion de A-module.

Tout d'abord, il est clair que B contient A, donc en particulier 0 et 1. Soient maintenant $\alpha, \beta \in B$. Par définition, il existe $P, Q \in A[X]$ unitaires tels que $P(\alpha) = 0$ et $Q(\beta) = 0$. Le polynôme $(-1)^n P(-X)$ est un polynôme unitaire à coefficients dans A, et annulant $-\alpha$. Ainsi, $-\alpha \in B$. Soit $\overline{L}$ une clôture algébrique de L, et écrivons maintenant

$$P = \prod_{i=1}^{n} (X - \alpha_i), \quad Q = \prod_{j=1}^{m} (X - \beta_j), \quad \alpha_i, \beta_j \in \overline{L},$$

avec $\alpha_1 = \alpha$ et $\beta_1 = \beta$. Définissons deux polynômes unitaires de $\overline{L}[X]$ par

$$R = \prod_{i=1}^{n} \prod_{j=1}^{m} (X - (\alpha_i + \beta_j)) \quad \text{et} \quad S = \prod_{i=1}^{n} \prod_{j=1}^{m} (X - \alpha_i \beta_j).$$

Les polynômes $\widetilde{R} = \prod_{j=1}^{m} (X - (Y + \beta_j))$ et $\widetilde{S} = \prod_{j=1}^{m} (X - \beta_j Y)$ sont des polynômes de $\overline{L}[X, Y]$ dont les coefficients sont des polynômes symétriques en $\beta_1, \ldots, \beta_m$ à coefficients dans A. Le théorème 1.9 montre alors que ces coefficients sont des polynômes en les coefficients de Q à coefficients dans A. En particulier, ce sont des éléments de A. Par conséquent, les polynômes $\widetilde{R}$, $\widetilde{S}$ sont à coefficients dans A. On remarque maintenant que l'on a les égalités

$$R = \prod_{i=1}^{n} \widetilde{R}(X, \alpha_i) \quad \text{et} \quad S = \prod_{i=1}^{n} \widetilde{S}(X, \alpha_i).$$

On en déduit facilement que les coefficients de R et S sont des polynômes symétriques à coefficients dans A en $\alpha_1, \ldots, \alpha_n$. Un raisonnement similaire au précédent implique que R, S sont des polynômes unitaires à coefficients dans A. Or, par construction, on a

$$R(\alpha + \beta) = 0 \quad \text{et} \quad S(\alpha\beta) = 0.$$

Ainsi, $\alpha + \beta$, $\alpha\beta \in B$, d'où le résultat.

On va maintenant introduire la notion d'anneau intégralement clos.

4.8. Définition. Soit A un anneau. On dit que A est *intégralement clos* s'il est intègre, et si les seuls éléments de K_A qui soient entiers sur A sont les éléments de A.

4.9. Exemples

(1) Tout anneau factoriel est intégralement clos.

En effet, soit $x = \dfrac{a}{b} \in K_A$. Après simplification par un pgcd de a et b, on peut supposer a et b premiers entre eux. Supposons que x soit entier sur A. Il existe donc $a_0, \ldots, a_{n-1} \in A$ tels que

$$x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0 = 0.$$

En multipliant cette équation par b^n, on obtient

$$a^n + a_{n-1}a^{n-1}b + \cdots + a_1 a b^{n-1} + a_0 b^n = 0.$$

Supposons maintenant que b possède un diviseur irréductible $\pi \in A$. Alors, π divise $a_{n-1}a^{n-1}b + \cdots + a_0 b^n$. En utilisant l'équation précédente, on voit que $\pi \mid a^n$. Puisque π est irréductible et que A est factoriel, on en déduit que $\pi \mid a$. Or, cela est impossible, car a et b sont premiers entre eux. Donc b est inversible dans A, et $x = ab^{-1} \in A$.

(2) L'anneau $\mathbb{Z}[\sqrt{5}]$ n'est pas intégralement clos.

En effet, $\phi = \dfrac{1 + \sqrt{5}}{2} \in K_A = \mathbb{Q}(\sqrt{5})$ est entier sur $\mathbb{Z}$, donc sur $\mathbb{Z}[\sqrt{5}]$, mais $\phi \notin \mathbb{Z}[\sqrt{5}]$.

On peut maintenant donner une caractérisation intéressante des éléments entiers sur A.

4.10. Théorème. *Soit A un anneau intégralement clos, et soit L/K_A une extension. Alors, $\alpha \in L$ est entier sur A si, et seulement si, son polynôme minimal sur K_A est à coefficients dans A.*

Démonstration. Soit P le polynôme minimal de α sur K_A. Si $P \in A[X]$, alors α est entier sur A, par définition. Inversement, supposons que $\alpha \in L$ soit entier sur A. Il existe donc un polynôme $Q \in A[X]$ unitaire non constant tel que $Q(\alpha) = 0$. Soient $\alpha = \alpha_1, \ldots, \alpha_n$ les racines (non nécessairement distinctes) de P dans une clôture algébrique $\overline{K}_A$ de K_A. Puisque l'on a $P \mid Q$ par définition du polynôme minimal de α, on obtient

$$Q(\alpha_i) = 0, \qquad \text{pour tout } i \in [\![1, n]\!],$$

et donc les éléments α_i sont entiers sur A. Comme les coefficients de P sont des sommes ou des différences de produits des α_i, ce sont des éléments entiers sur A par le corollaire 4.6. Ainsi, les coefficients de P sont des éléments de K_A entiers sur A. Puisque A est intégralement clos, ils appartiennent à A, et donc $P \in A[X]$. □

4.11. Remarque. Ce critère est surtout utile pour démontrer qu'un élément algébrique sur K_A n'est pas entier sur A.

Par exemple, $\dfrac{\sqrt{5}}{2} \in \mathbb{Q}(\sqrt{5})$ n'est pas entier sur $\mathbb{Z}$, car son polynôme minimal sur $\mathbb{Q}$ est $X^2 - \dfrac{5}{4}$.

En revanche, pour montrer qu'un élément $\alpha \in \overline{K}_A$ est entier sur A, il n'est pas nécessaire de calculer son polynôme minimal. Il suffit pour cela d'exhiber un polynôme unitaire de $A[X]$ annulant α.

On finit ces considérations par un théorème qui répond à notre interrogation première.

4.12. Théorème. *Soit A un anneau intègre. Alors, tout polynôme unitaire irréductible de $A[X]$ est encore irréductible dans l'anneau $K_A[X]$ si, et seulement si, l'anneau A est intégralement clos.*

Démonstration. Supposons que tout polynôme unitaire irréductible de $A[X]$ soit encore irréductible dans $K_A[X]$ et soit $\alpha \in K_A$ un élément entier sur A. Il existe donc $P \in A[X]$ unitaire tel que $P(\alpha) = 0$. Choisissons P de degré minimal. Alors, P est irréductible dans $A[X]$. En effet, P est non nul, et possède au moins une racine dans K_A, donc P est non constant. Comme A

est intègre, on a $A[X]^\times = A^\times$. En particulier, P est non inversible. Supposons maintenant que $P = QR$, avec $Q, R \in A[X]$. Comme P est unitaire, en comparant les coefficients dominants, on voit que les coefficients dominants de P et Q sont inversibles, et inverses l'un de l'autre. Quitte à multiplier Q et R par un élément inversible, on peut les supposer unitaires. Par intégrité de A, on a $Q(\alpha) = 0$ ou $R(\alpha) = 0$, disons $Q(\alpha) = 0$. Mais alors, R est constant, car sinon on aurait $\deg(Q) < \deg(P)$, ce qui contredirait le choix de P. Comme R est unitaire, $R = 1 \in A[X]^\times$, d'où l'irréductibilité de P dans $A[X]$. Par hypothèse, P est irréductible dans $K_A[X]$. C'est donc le polynôme minimal de α sur K_A, à savoir $X - \alpha$. Comme P est à coefficients dans A, on a $\alpha \in A$, et A est intégralement clos.

Réciproquement, supposons A intégralement clos, et soit $P \in A[X]$ un polynôme unitaire irréductible non constant. Soit $Q \in K_A[X]$ un facteur irréductible unitaire de P, et soit $\alpha \in \overline{K}_A$ une racine de Q. Alors, α est entier sur A, puisque $P(\alpha) = 0$. Par définition, Q est le polynôme minimal de α sur K_A. D'après le théorème 4.10, on a $Q \in A[X]$. Puisque Q est unitaire, Q divise alors P dans $A[X]$. De plus, Q est irréductible dans l'anneau $A[X]$. En effet, Q est un polynôme non constant, car il est irréductible dans $K_A[X]$. De plus, si Q est un produit de deux polynômes à coefficients dans A, on peut comme précédemment les supposer unitaires. En regardant cette décomposition dans $K_A[X]$, on constate qu'un des deux polynômes est constant, et par conséquent égal à 1. Comme $Q \mid P$, et puisque P et Q sont tous deux unitaires et irréductibles dans $A[X]$, on obtient finalement que $P = Q$. En particulier, P est irréductible dans l'anneau $K_A[X]$, ce qui achève la démonstration. □

4.13. Remarque. En particulier, lorsque A est intégralement clos, les théorèmes d'irréductibilité énoncés dans le chapitre XXI peuvent être utilisés pour établir l'irréductibilité de polynômes de $K_A[X]$.

5. Exercices

E1. Montrer qu'un corps algébriquement clos est infini.

E2. Soit K un corps, et soit $f \in K[X]$ un polynôme de degré $n \geqslant 1$. Enfin, soit L/K une extension de degré $m \geqslant 1$ premier à n. Montrer que f est irréductible dans l'anneau $K[X]$ si, et seulement si, f est irréductible dans l'anneau $L[X]$.

Indication. Considérer la sous-extension de $\overline{K}$ engendrée par une racine de f.

E3. Soit $K = \mathbb{F}_2(X_1^2, X_2^2)$ et soit $L = K(X_1, X_2)$.

a. Montrer que $[L : K] = 4$.

b. Montrer que pour tout $\alpha \in L$, on a $\alpha^2 \in K$, puis que $[K(\alpha) : K] \leqslant 2$.

c. En déduire que L/K n'est pas monogène.

d. Montrer que L/K est le corps des racines d'un polynôme de $K[X]$.

e. Montrer que L/K n'est pas le corps des racines d'un polynôme irréductible de $K[X]$.

E4. Soit K un corps.

a.

(i) Soit $d \in K^\times$. Montrer que chaque racine du polynôme $X^2 - d$ dans $\overline{K}$ engendre une même extension.

On note cette extension $K(\sqrt{d})/K$.

(ii) Montrer que $[K(\sqrt{d}) : K] \leqslant 2$, et que l'égalité a lieu si, et seulement si, on a $d \in K \setminus K^{\times 2}$.

(iii) Si d_1, $d_2 \in K^\times$, montrer que

$$K(\sqrt{d_1}) = K(\sqrt{d_2}) \iff d_2 d_1^{-1} \in K^{\times 2}.$$

b. On suppose que K est un corps de caractéristique 2.

(i) Montrer que $\wp(K) = \{x^2 - x \mid x \in K\}$ est un sous-groupe additif de K.

(ii) Soit $a \in K$. Montrer que chaque racine du polynôme $X^2 - X - a$ dans $\overline{K}$ engendre une même extension.

On note $K\big(\wp^{-1}(a)\big)/K$ cette extension.

(iii) Montrer que $[K\big(\wp^{-1}(a)\big) : K] \leqslant 2$, avec égalité si, et seulement si, on a $a \in K \setminus \wp(K)$.

(iv) Si a_1, $a_2 \in K$, montrer que

$$K\big(\wp^{-1}(a_1)\big) = K\big(\wp^{-1}(a_2)\big) \iff a_2 - a_1 \in \wp(K).$$

(v) Soit $d \in K^\times \setminus K^{\times 2}$ et soit $a \in K \setminus \wp(K)$. Montrer que $K(\sqrt{d})/K$ et $K\big(\wp^{-1}(a)\big)/K$ ne sont pas isomorphes.

c.

(i) On suppose que $\mathrm{car}(K) \neq 2$. Montrer que toute extension L/K de degré 2 est de la forme

$$K(\sqrt{d})/K, \quad d \in K^\times \setminus K^{\times 2}.$$

(ii) On suppose que $\mathrm{car}(K) = 2$. Montrer que toute extension L/K de degré 2 est de la forme

$$K(\sqrt{d})/K,\ d \in K^\times \setminus K^{\times 2} \quad \text{ou} \quad K\big(\wp^{-1}(a)\big)/K,\ a \in K \setminus \wp(K).$$

E5. Soit p un nombre premier. Montrer que $\bigcup_{n\geqslant 1} \mathbb{F}_{p^{n!}}$ est une clôture algébrique de $\mathbb{F}_p$.

Indication. On pourra utiliser l'exercice XXI-10.

E6.

a. Soit K un corps, soit $a \in K$ et soit p un nombre premier. Montrer que $X^p - a$ est irréductible dans $K[X]$ si, et seulement si, il n'a pas de racines dans K.

Indication. Si $X^p - a = PQ$ avec $P, Q \in K[X]$ non constants, montrer en décomposant $X^p - a$ en facteurs de degré 1 dans $\overline{K}[X]$, que l'on a les égalités

$$\big((-1)^n P(0)\big)^p = a^n \quad \text{et} \quad \big((-1)^{p-n} Q(0)\big)^p = a^{p-n},$$

avec $\deg(P) = n$. En utilisant une identité de Bézout, montrer alors que a est une puissance p-ième dans K.

b. Soient p, ℓ deux nombres premiers impairs.
On suppose que $\ell \equiv -1 \mod 3$ et que $\overline{p}$ est un générateur de $\mathbb{F}_\ell^\times$. En réduisant modulo p et modulo 2, montrer que $X^{\ell+1} - X + p$ est irréductible dans $\mathbb{Z}[X]$.

Indication. Utiliser la question précédente.

E7. Soit K un corps de caractéristique p, avec p premier, et soit $a \in K$. Montrer que $P = X^p - X - a$ est irréductible dans $K[X]$ si, et seulement si, P ne possède pas de racines dans K.

Indication. Si $\alpha \in \overline{K}$, quelles sont les autres racines de P ? Si $Q \in K[X]$ est un diviseur de P de degré $d \in [\![1, p-1]\!]$, regarder le terme en X^{d-1}.

E8. Retrouver l'unicité d'un corps des racines d'un polynôme en utilisant le théorème de Steinitz.

E9. Soit L/K une extension algébrique. Comparer les clôtures algébriques de K et de L.

E10. Soit p un nombre premier, soit $q \geqslant 2$ une puissance de p et soit $n \geqslant 1$ un entier tel que $p \nmid n$. On note $\overline{\Phi}_n$ la réduction de Φ_n modulo p. Enfin, soit $r = o(\overline{q}) \in (\mathbb{Z}/n\mathbb{Z})^\times$.

On se fixe $\pi \in \mathbb{F}_q[X]$ un facteur irréductible unitaire de degré $s \geqslant 1$ de $\overline{\Phi}_n$, ainsi qu'un corps de rupture $K = \mathbb{F}_q(\alpha)$ de π.

a. Montrer par l'absurde que $X^n - 1 \in \mathbb{F}_q[X]$ est un produit de polynômes irréductibles unitaire deux à deux distincts. Montrer que les racines du polynôme $\overline{\Phi}_n$ dans $\mathbb{F}_q$ sont exactement les éléments de $\mathbb{F}_q^\times$ d'ordre n.

b. Montrer que $\alpha^{q^s-1} = 1$. En déduire que $n \mid q^s - 1$, puis $s \geqslant r$.

c. Montrer que $\alpha^{q^r} = \alpha$, puis que $s = r$.

d. En déduire que $\overline{\Phi}_n$ se décompose dans $\mathbb{F}_q[X]$ en produit de $\dfrac{\varphi(n)}{r}$ polynômes irréductibles unitaires de degré r.

e. Montrer l'équivalence des trois propriétés suivantes :

(1) Il existe un nombre premier p ne divisant pas n tel que Φ_n soit irréductible modulo p ;

(2) le groupe $(\mathbb{Z}/n\mathbb{Z})^\times$ est cyclique ;

(3) on a $n = 1$, 2, 4, ℓ^m ou $2\ell^m$, où $m \geqslant 1$ et ℓ est un nombre premier impair.

On pourra utiliser librement le fait difficile suivant, dû à Dirichlet : pour tout entier $a \in \mathbb{Z}$ premier à n, il existe un nombre premier p tel que $p \equiv a\ [n]$ (il y en a même une infinité, cf. [43] pour une démonstration).

E11. Si B est un anneau et A est un sous-anneau de B, on dit que B est entier sur A si tout élément de B est entier sur A.

a. Soit C un anneau, et soient A, B des sous-anneaux de C tels que

$$A \subset B \subset C.$$

Montrer que C est entier sur A si, et seulement si, B est entier sur A et C est entier sur B.

b. Soit A un anneau intègre, et soit L/K_A une extension. Soit B le sous-anneau des éléments de L entiers sur A. Montrer que B est intégralement clos.

E12. Soit A un anneau intègre non intégralement clos. On se propose d'exhiber un idéal de A non principal. On note B le sous-anneau des éléments de K_A qui sont entiers sur A. Soit $x = \dfrac{a}{s} \in B \setminus A$ un élément de K_A entier sur A qui n'est pas dans A, et soit $I = (a, s)$.
On note IB l'idéal de B engendré par les éléments de I.

a. Montrer que $IB = sB$, puis en déduire que si $I = \alpha A$, alors $\alpha = us$, avec $u \in B^\times$.

b. Montrer alors l'existence de $v \in A$ tel que $uv = 1$.

c. En exploitant le fait que u est entier sur A, montrer l'existence d'un polynôme $Q \in A[X]$ tel que $1 + vQ(v) = 0$.

En déduire alors que $u \in A$, puis que $u \in A^\times$.

d. Conclure.

E13. Soit K un corps, et soit $\mathscr{A}$ une K-algèbre associative, commutative et unitaire.

On dit qu'un groupe G agit sur $\mathscr{A}$ par automorphismes d'algèbre, si G agit sur $\mathscr{A}$ et son action vérifie les propriétés suivantes :

(1) pour tout $g \in G$, $g{\cdot}1_{\mathscr{A}} = 1_{\mathscr{A}}$;

(2) pour tout $g \in G$, tout $x \in \mathscr{A}$, et tout $\lambda \in K$, on a $g{\cdot}(\lambda x) = \lambda(g{\cdot}x)$;

(3) pour tout $g \in G$, et tous $x_1, x_2 \in \mathscr{A}$, on a $g{\cdot}(x_1{+}x_2) = g{\cdot}x_1{+}g{\cdot}x_2$;

(4) pour tout $g \in G$, et tous $x_1, x_2 \in \mathscr{A}$, on a $g{\cdot}(x_1x_2) = (g{\cdot}x_1)(g{\cdot}x_2)$.

a. Montrer que si G agit sur $\mathscr{A}$ par automorphismes d'algèbre, l'ensemble $\mathscr{A}^G$ est une sous-algèbre de $\mathscr{A}$.

On suppose dorénavant que G est d'ordre $n \geqslant 1$, et que $\mathscr{A}$ est de type fini, i.e. que $\mathscr{A} = K[\alpha_1, \ldots, \alpha_r]$, où $\alpha_1, \ldots, \alpha_r \in \mathscr{A}$. On se propose alors de démontrer que $\mathscr{A}^G$ est aussi de type fini [5].

b. Si $\alpha \in \mathscr{A}$, montrer que $f_\alpha = \prod_{g \in G} (T - g{\cdot}\alpha) \in \mathscr{A}[T]$ est un polynôme unitaire de degré n à coefficients dans $\mathscr{A}^G$.

Pour tout $i \in [\![1, r]\!]$, on note $f_{\alpha_i} = T^n + a_{i,n-1}T^{n-1} + \cdots + a_{i,1}T + a_{i,0}$, avec $a_{i,j} \in \mathscr{A}^G$, et on pose

$$\mathscr{B} = K[a_{1,0}, \ldots, a_{1,n-1}, \ldots, a_{n,0}, \ldots, a_{r,n-1}] \subset \mathscr{A}^G.$$

c. Montrer que $\alpha_1, \ldots, \alpha_r \in \mathscr{A}$ sont entiers sur $\mathscr{B}$, et en déduire que $\mathscr{A}$ est un $\mathscr{B}$-module de type fini.

d. Justifier que $\mathscr{A}$ est un $\mathscr{B}$-module noethérien. En déduire que $\mathscr{A}^G$ est un $\mathscr{B}$-module de type fini, et conclure.

Indication. Utiliser l'exercice XX-5.

5. Ce résultat est dû à Noether.

On finit cet exercice en donnant une famille d'exemples.

Pour tout $P \in K[X_1, \ldots, X_n]$, et tout $M \in \mathrm{GL}_n(K)$, on pose

$$M \bullet P = P\Big(\sum_{j=1}^{n} a_{1j}X_j, \ldots, \sum_{j=1}^{n} a_{nj}X_j\Big),$$

où $M^{-1} = (a_{ij})_{i,j}$.

e. Montrer que l'application

$$\begin{aligned} \mathrm{GL}_n(K) \times K[X_1, \ldots, X_n] &\longrightarrow K[X_1, \ldots, X_n] \\ (g, P) &\longmapsto M \bullet P \end{aligned}$$

définit une action de $\mathrm{GL}_n(K)$ sur $K[X_1, \ldots, X_n]$ par automorphismes d'algèbre.

f. Si $\rho : G \longrightarrow \mathrm{GL}_n(K)$ est un morphisme de groupes, montrer que l'application

$$\begin{aligned} G \times K[X_1, \ldots, X_n] &\longrightarrow K[X_1, \ldots, X_n] \\ (g, P) &\longmapsto g \cdot P = \rho(g) \bullet P \end{aligned}$$

définit une action de G sur $K[X_1, \ldots, X_n]$ par automorphismes d'algèbre.

En déduire que si G est fini, $K[X_1, \ldots, X_n]^G$ est une K-algèbre associative, commutative et unitaire de type fini.

g. Expliciter l'action de $\mathfrak{S}_n$ sur la K-algèbre $K[X_1, \ldots, X_n]$ lorsque le morphisme $\rho : \mathfrak{S}_n \longrightarrow \mathrm{GL}_n(K)$ est défini par

$$\rho(\tau) = M_\tau = \big(\delta_{i,\tau(j)}\big)_{i,j}, \qquad \text{pour tout } \tau \in \mathfrak{S}_n.$$

Décrire $K[X_1, \ldots, X_n]^{\mathfrak{S}_n}$.

h. Soit $\rho : \mathbb{Z}/2\mathbb{Z} \longrightarrow \mathrm{GL}_2(\mathbb{F}_2)$ l'unique morphisme de groupes défini par

$$\rho(\overline{1}) = \begin{pmatrix} \overline{1} & \overline{1} \\ \overline{0} & \overline{1} \end{pmatrix}.$$

Montrer que $\mathbb{F}_2[X, Y]^{\mathbb{Z}/2\mathbb{Z}} = \mathbb{F}_2[Y, X(X+Y)]$.

Indication. Vérifier que tout $P \in \mathbb{F}_2[X, Y]$ s'écrit de manière unique sous la forme

$$P = \sum_{n,m \geqslant 0} a_{n,m} X^n (X+Y)^m,$$

où les éléments $a_{n,m} \in \mathbb{F}_2$ sont presque tous nuls.

En déduire que si $P \in \mathbb{F}_2[X, Y]^{\mathbb{Z}/2\mathbb{Z}}$, alors P est un polynôme symétrique en X et $X+Y$ à coefficients dans $\mathbb{F}_2$.

Conclure en remarquant que X et $X+Y$ sont les racines de

$$T^2 - YT + X(X+Y) \in \mathbb{F}_2(X,Y)[T].$$

E14. Si K est un sous-corps de $\mathbb{C}$, on note $\mathcal{O}_K$ l'ensemble des éléments de K qui sont entiers sur $\mathbb{Z}$. C'est un sous-anneau de $\mathbb{C}$ contenant $\mathbb{Z}$.

a. Montrer que $\mathcal{O}_{\mathbb{Q}} = \mathbb{Z}$.

b. Montrer que $\mathcal{O}_{\mathbb{C}}$ est un anneau intègre qui n'est pas un corps, et qui ne possède aucun élément irréductible.

Indication. On pourra montrer que si, $x \in \mathcal{O}_{\mathbb{C}}$, alors $\sqrt{x} \in \mathcal{O}_{\mathbb{C}}$.

c. Montrer que $\mathcal{O}_{\mathbb{C}}$ n'est pas noethérien.

d. Montrer que pour tout élément $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$, on a $\sigma(\mathcal{O}_K) \subset \mathcal{O}_K$.

e. Soit $d \in \mathbb{Z}$ un entier sans facteurs carrés, et soit

$$K = \mathbb{Q}(\sqrt{d}) = \{a + b\sqrt{d} \mid a, b \in \mathbb{Z}\}.$$

Montrer $\mathcal{O}_K = \mathbb{Z}[\sqrt{d}]$ si $d \not\equiv 1$ [4] et que $\mathcal{O}_K = \mathbb{Z}\left[\frac{1+\sqrt{d}}{2}\right]$ si $d \equiv 1$ [4].

Indication. Montrer que si $a + b\sqrt{d} \in \mathcal{O}_K$, alors $a - b\sqrt{d} \in \mathcal{O}_K$, puis en déduire que $2a, a^2 - db^2, 2bd \in \mathbb{Z}$. On pourra utiliser, après avoir justifié son existence, un élément $\sigma \in \mathrm{Gal}(K/\mathbb{Q})$ vérifiant

$$\sigma(\sqrt{d}) = -\sqrt{d}.$$

Chapitre XXVII

Plongements et extensions séparables

Dans tout ce qui suit, on se fixe un corps de base K, et une clôture algébrique $\overline{K}/K$ de K. Toutes les extensions algébriques considérées (en particulier les extensions de degré fini) seront alors des sous-extensions de $\overline{K}/K$.

Un des buts principaux de ce chapitre est de comprendre un peu mieux le groupe $\mathrm{Gal}(L/K)$ lorsque L/K est une extension de degré fini. En particulier, on aimerait étudier son ordre. Lorsque $L = K(\alpha)$, la réponse est donnée par le théorème XXVI-2.2 (appliqué à $\iota = \mathrm{Id}_K$) : l'ordre de $\mathrm{Gal}(L/K)$ est le nombre de racines de $\mu_{\alpha,K}$ dans L.

Lorsque L est engendrée par plus d'un générateur, la réponse paraît moins immédiate. Par exemple, supposons que $L = K(\alpha_1, \alpha_2)$. Comme on a la tour d'extensions

$$K \subset K(\alpha_1) \subset K(\alpha_1)(\alpha_2) = L,$$

il est tentant d'essayer de se ramener au cas précédent. Hélas, ce n'est pas si simple : si $\sigma \in \mathrm{Gal}(L/K)$, on obtient par restriction un morphisme de K-algèbres $\sigma' : K(\alpha_1) \longrightarrow L$, mais qui n'est a priori pas un automorphisme, car il n'y aucune raison pour que son image soit incluse dans $K(\alpha_1)$. En revanche, σ' peut être vu comme un morphisme d'extensions $K(\alpha_1) \longrightarrow \overline{K}$.

Inversement, un tel morphisme s'étend en un morphisme $L \longrightarrow \overline{K}$ par le théorème XXVI-2.2. Ainsi, les bons objets à étudier, à la fois d'un point de vue théorique et d'un point de vue calculatoire, sont les morphismes d'extensions $L \longrightarrow \overline{K}$. Remarquons qu'en particulier l'ensemble de ces morphismes contient $\mathrm{Gal}(L/K)$ (puisque L est un sous-corps de $\overline{K}$).

Dans ce chapitre, nous allons nous intéresser à l'ensemble des morphismes de K-algèbres $L \longrightarrow \overline{K}$.

1. Plongements

Commençons par définir une des notions principales de ce paragraphe.

1.1. Définition. Soit L/K une extension de corps algébrique. Un *plongement* de L/K est un morphisme d'extensions $\sigma : L \longrightarrow \overline{K}$. L'ensemble des plongements de L/K est noté $\mathscr{X}(L/K)$.

Le lemme suivant, bien que de démonstration fort simple, est très important pour la suite. Comme nous en aurons besoin en toute généralité, commençons par introduire la notion de monoïde et de morphismes de monoïdes.

1.2. Définition. Un *monoïde* est un ensemble M non vide muni d'une loi interne $*$ associative et possédant un élément neutre 1_M (nécessairement unique d'après le lemme I-3.5).

Un *morphisme de monoïdes* est une application $f : (M, *) \longrightarrow (M', \perp)$ entre deux monoïdes vérifiant :

(i) pour tous $x, y \in M$, $f(x * y) = f(x) \perp f(y)$;
(ii) $f(1_M) = 1_{M'}$.

1.3. Exemples

(1) Tout groupe est en particulier un monoïde, et tout morphisme de groupes $f : G \longrightarrow G'$ est un morphisme de monoïdes [1].

(2) Si A est un anneau, $(A, \cdot)$ est un monoïde qui n'est pas un groupe (sauf si A est trivial), et tout morphisme d'anneaux $f : A \longrightarrow B$ est un morphisme de monoïdes $(A, \cdot) \longrightarrow (B, \cdot)$.

(3) Les couples $(\mathbb{N}, +)$ et $(\mathbb{N}, \cdot)$ sont des monoïdes.

On peut maintenant énoncer le lemme d'indépendance de Dedekind.

1.4. Lemme (d'indépendance de Dedekind). *Soit $(M, *)$ un monoïde et soit E un corps.*

Alors, toute famille de morphismes de monoïdes $(M, *) \longrightarrow (E, \cdot)$ deux à deux distincts est E-libre dans le E-espace vectoriel $\mathscr{F}(M, E)$.

Démonstration. Il suffit bien entendu de se réduire aux familles finies.

Montrons par récurrence que pour tout entier $n \geqslant 1$, n morphismes de monoïdes $(M, *) \longrightarrow (E, \cdot)$ deux à deux distincts sont linéairement indépendants sur E.

1. Puisque $f(1_G) = 1_{G'}$.

Si $n = 1$, c'est clair puisqu'un morphisme $(M, *) \longrightarrow (E, \cdot)$ est non nul (il envoie 1_M sur $1_E \neq 0_E$).

Soit $n \geqslant 1$, et supposons la propriété soit vraie pour n morphismes de monoïdes. Soient $\varphi_1, \ldots, \varphi_{n+1} : (M, *) \longrightarrow (E, \cdot)$, $n+1$ morphismes de monoïdes, deux à deux distincts, et soient $\lambda_1, \ldots, \lambda_{n+1} \in E$ tels que

$$\lambda_1\varphi_1 + \cdots + \lambda_{n+1}\varphi_{n+1} = 0.$$

Fixons $x \in M$. On a alors

$$\lambda_1\varphi_1(x * y) + \cdots + \lambda_{n+1}\varphi_{n+1}(x * y) = 0, \quad \text{pour tout } y \in M,$$

soit encore

$$\lambda_1\varphi_1(x)\varphi_1(y) + \cdots + \lambda_{n+1}\varphi_{n+1}(x)\varphi_{n+1}(y) = 0, \quad \text{pour tout } y \in M.$$

Mais on a aussi

$$\lambda_1\varphi_1(y) + \cdots + \lambda_{n+1}\varphi_{n+1}(y) = 0, \quad \text{pour tout } y \in M.$$

En multipliant cette dernière équation par $\varphi_{n+1}(x)$ et en soustrayant à la précédente, on obtient

$$\lambda_1\big(\varphi_1(x) - \varphi_{n+1}(x)\big)\varphi_1(y) + \cdots + \lambda_n\big(\varphi_n(x) - \varphi_{n+1}(x)\big)\varphi_n(y) = 0,$$

pour tout $y \in M$. Autrement dit,

$$\lambda_1\big(\varphi_1(x) - \varphi_{n+1}(x)\big)\varphi_1 + \cdots + \lambda_n\big(\varphi_n(x) - \varphi_{n+1}(x)\big)\varphi_n = 0.$$

Par hypothèse de récurrence, on a alors

$$\lambda_i\big(\varphi_i(x) - \varphi_{n+1}(x)\big) = 0, \quad \text{pour tout } i \in [\![1, n]\!].$$

Cela étant vrai pour tout $x \in M$, cela se récrit

$$\lambda_i(\varphi_i - \varphi_{n+1}) = 0, \quad \text{pour tout } i \in [\![1, n]\!].$$

Les morphismes $\varphi_1, \ldots, \varphi_{n+1}$ étant deux à deux distincts, on obtient

$$\lambda_i = 0, \quad \text{pour tout } i \in [\![1, n]\!].$$

Par conséquent, $\lambda_{n+1}\varphi_{n+1} = 0$, d'où également $\lambda_{n+1} = 0$ comme on l'a déjà vu précédemment.

Cela achève la récurrence, ainsi que la démonstration. □

1.5. Corollaire. *Soient L et M deux corps.*

Alors, toute famille de morphismes d'anneaux $L \longrightarrow M$ deux à deux distincts est M-libre dans le M-espace vectoriel $\mathscr{F}(L, M)$.

En particulier, pour toute extension L/K, les éléments de $\mathscr{X}(L/K)$ (resp. les éléments de $\mathrm{Gal}(L/K)$) forment une famille $\overline{K}$-libre de $\mathscr{F}(L, \overline{K})$ (resp. une famille L-libre de $\mathscr{F}(L, L)$).

Démonstration. Il suffit de remarquer qu'une famille de morphismes d'anneaux $L \longrightarrow M$ deux à deux distincts induit une famille de morphismes de monoïdes $(L,\cdot) \longrightarrow (M,\cdot)$ deux à deux distincts, et d'appliquer le lemme précédent. □

On va maintenant étudier de plus près les plongements d'une extension monogène.

1.6. Lemme. *Soit $\alpha \in \overline{K}$, et soit $L = K(\alpha)$.*

Alors, tout morphisme d'anneaux $\iota : K \longrightarrow \overline{K}$ possède au plus $[L : K]$ prolongements $\sigma : L \longrightarrow \overline{K}$ à L. En particulier, on a

$$|\mathscr{X}(L/K)| \leqslant [L : K].$$

De plus, les propriétés suivantes sont équivalentes :

(1) *tout morphisme d'anneaux $\iota : K \longrightarrow \overline{K}$ admet exactement $[L : K]$ prolongements à L ;*
(2) *on a l'égalité $|\mathscr{X}(L/K)| = [L : K]$;*
(3) *le polynôme $\mu_{\alpha,K}$ ne possède que des racines simples dans $\overline{K}$.*

Démonstration. Gardons les notations de l'énoncé.

Par le théorème XXVI-2.2, ι possède autant de prolongements à L que le polynôme $\iota(\mu_{\alpha,K})$ possède de racines distinctes dans $\overline{K}$. En particulier, il y a au plus $\deg\big(\iota(\mu_{\alpha,K})\big)$ prolongements. Or, on a

$$\deg\big(\iota(\mu_{\alpha,K})\big) = \deg(\mu_{\alpha,K}) = [L : K],$$

d'où la première partie.

Un plongement de L/K n'étant rien d'autre qu'un prolongement de l'inclusion $K \subset \overline{K}$, on en déduit que $|\mathscr{X}(L/K)|$ est le nombre de racines distinctes de $\mu_{\alpha,K}$ dans $\overline{K}$. On obtient alors l'inégalité

$$|\mathscr{X}(L/K)| \leqslant [L : K],$$

ainsi que la chaîne d'implications « (1) $\Longrightarrow$ (2) $\Longrightarrow$ (3) ».

Il reste à démontrer « (3) $\Longrightarrow$ (1) ». Remarquons tout d'abord que pour tout $\iota : K \longrightarrow \overline{K}$, et tout $f \in K[X]$, on a

$$\operatorname{disc}\big(\iota(f)\big) = \iota\big(\operatorname{disc}(f)\big),$$

où $\operatorname{disc}(P)$ désigne le discriminant de $P \in K[X]$ (cf. définition XXVI-1.12). En effet, par la proposition XXVI-1.11, il existe $h \in A_0[Y_0, \ldots, Y_n]$ tel que

$$\operatorname{disc}(f) = h(a_0, \ldots, a_n),$$

où $a_0, \ldots, a_n \in K$ sont les coefficients de f et A_0 est le sous-anneau engendré par 1_K.

Or, ι étant un morphisme d'anneaux, il fixe les éléments de A_0. On a alors

$$\iota\big(\mathrm{disc}(f)\big) = \iota\big(h(a_0, \ldots, a_n)\big) = h\big(\iota(a_0), \ldots, \iota(a_n)\big) = \mathrm{disc}\big(\iota(f)\big).$$

Le morphisme $\iota : K \longrightarrow \overline{K}$ étant injectif, on en déduit que $f \in K[X]$ n'a que des racines simples dans $\overline{K}$ si, et seulement si, $\iota(f) \in \overline{K}[X]$ ne possède que des racines simples dans $\overline{K}$, par la proposition XXVI-1.11 [2].

Supposons maintenant que $\mu_{\alpha,K} \in K[X]$ ne possède que des racines simples dans $\overline{K}$ et soit $\iota : K \longrightarrow \overline{K}$ un morphisme d'anneaux. D'après ce qui précède, $\iota(\mu_{\alpha,K})$ ne possède que des racines simples dans $\overline{K}$. Les considérations faites en début de démonstration montrent alors que ι possède $[L : K]$ prolongements à L, ce qui achève la démonstration. □

Ce résultat nous conduit naturellement à la définition suivante.

1.7. Définition. On dit qu'un polynôme $P \in K[X]$ non constant est *séparable sur* K si P n'a que des racines simples dans $\overline{K}$. On dit que $\alpha \in \overline{K}$ est *séparable sur* K si $\mu_{\alpha,K}$ est séparable sur K. On dit qu'une extension L/K est *séparable* si L/K est algébrique, et si tout élément de L est séparable sur K.

1.8. Remarque. Soit $P \in K[X]$ un polynôme non constant. Alors, les propriétés suivantes sont équivalentes :

(1) le polynôme P est séparable ;

(2) il existe une extension L/K dans laquelle P est scindé à racines simples ;

(3) pour toute extension L/K dans laquelle P est scindé, P est à racines simples.

Les implications « (1) $\Longrightarrow$ (2) » et « (3) $\Longrightarrow$ (1) » sont évidentes. Supposons maintenant qu'il existe une extension L/K dans laquelle P est scindé à racines simples, et soit L'/K une extension dans laquelle P est scindé. Alors, L/K et L'/K contiennent des corps des racines E/K et E'/K de P. De plus, P est scindé à racines simples dans E. Ces corps des racines sont isomorphes par la proposition XXVI-2.7. Soit $\varphi : E \xrightarrow{\sim} E'$ un tel isomorphisme. On a alors

$$\varphi\big(P(\alpha)\big) = P\big(\varphi(\alpha)\big).$$

Comme φ est bijective, on en déduit aisément que φ induit une bijection entre l'ensemble des racines de P dans E et l'ensemble des racines de P dans E'. En particulier, P n'a que des racines simples dans E', donc dans L', d'où le résultat.

2. Il y a une façon moins élaborée de montrer ce fait, cf. remarque 3.3.

En particulier, la définition de séparabilité ne dépend pas de la clôture algébrique choisie.

Cela démontre également qu'un élément $\alpha \in L/K$ est séparable sur K si, et seulement si, il existe une extension de K (non nécessairement algébrique) dans laquelle $\mu_{\alpha,K}$ est scindé à racines simples. Ce fait sera utilisé dans la suite sans référence ultérieure.

1.9. Exemples

(1) Tout élément de $\alpha \in K$ est séparable sur K puisque $\mu_{\alpha,K} = X - \alpha$.

(2) Le nombre complexe i est séparable sur $\mathbb{Q}$, puisque

$$\mu_{i,\mathbb{Q}} = X^2 + 1 = (X - i)(X + i).$$

(3) Soit $K = \mathbb{F}_2(T^2)$ et $L = \mathbb{F}_2(T)$. Alors, T n'est pas séparable sur K, puisque

$$\mu_{T,K} = X^2 - T^2 = (X - T)^2.$$

1.10. Remarque. Soit L/K une extension, et soit M/K une sous-extension. Si $\alpha \in L$ est séparable sur K, il est séparable sur M.

En effet, si $\alpha \in L$ est algébrique, alors $\mu_{\alpha,K}$ est un polynôme unitaire annulant α, que l'on peut voir à coefficients dans M. Ainsi, $\mu_{\alpha,M} \mid \mu_{\alpha,K}$. Par conséquent, si $\mu_{\alpha,K}$ n'a que des racines simples dans $\overline{K}$, il en est de même de $\mu_{\alpha,M}$. Ainsi, si L/K est séparable, L/M et M/K sont séparables (le fait que M/K soit séparable est clair).

On a alors le théorème suivant.

1.11. Théorème. *Soit L/K une extension de degré fini.*

Alors, tout morphisme d'anneaux $\iota : K \longrightarrow \overline{K}$ admet au plus $[L : K]$ prolongements $\sigma : L \longrightarrow \overline{K}$ à L. En particulier, on a

$$|\mathscr{X}(L/K)| \leqslant [L : K].$$

De plus, les propriétés suivantes sont équivalentes :

(1) *tout morphisme d'anneaux $\iota : K \longrightarrow \overline{K}$ admet exactement $[L : K]$ prolongements à L ;*

(2) *on a l'égalité $|\mathscr{X}(L/K)| = [L : K]$;*

(3) *l'extension L/K est séparable.*

Démonstration. Montrons le résultat par récurrence sur $[L : K]$. Plus précisément, pour tout $n \geqslant 1$, soit (H_n) la propriété suivante :

(H_n) Pour tout corps K, et toute extension L/K de degré $\leqslant n$, l'énoncé du théorème est vrai.

Si $n = 1$, alors $L = K$ et toutes les propriétés de l'énoncé sont trivialement vraies. Soit $n \geqslant 1$ un entier, et supposons que la propriété (H_n) soit vraie. Montrons que (H_{n+1}) est vraie. Soit K un corps, et soit L/K une extension de degré $\leqslant n+1$. On peut toujours supposer que $[L : K] = n+1 \geqslant 2$, car sinon il suffit d'appliquer (H_n). En particulier, $L \neq K$. Soit $\alpha \in L \setminus K$. On a donc $K(\alpha) \neq K$, et par conséquent $[L : K(\alpha)] \leqslant n$.

Remarquons maintenant que tout prolongement de ι à L s'obtient par extension d'un prolongement de ι à $K(\alpha)$. En effet, si $\sigma : L \longrightarrow \overline{K}$ est un prolongement de ι à L, alors $\sigma_{|K(\alpha)}$ est un prolongement de ι à $K(\alpha)$. Inversement, tout prolongement de ι à $K(\alpha)$ possède au plus $[L : K(\alpha)]$ extensions à L par hypothèse de récurrence, donc au moins une. De plus, il y a exactement n_α prolongements de ι à $K(\alpha)$ par le lemme 1.6, où n_α est le nombre de racines de $\iota(\mu_{\alpha,K})$ dans $\overline{K}$. On en déduit que ι admet au plus $n_\alpha[L : K(\alpha)]$ extensions à L. Comme $n_\alpha \leqslant [K(\alpha) : K]$ et que $[K(\alpha) : K][L : K(\alpha)] = [L : K]$, cela démontre le premier point.

Notons que tout cela est valable pour un élément $\alpha \in L \setminus K$ arbitraire.

Montrons maintenant l'équivalence des trois propriétés (toujours sous l'hypothèse (H_n)).

(1) $\Longrightarrow$ (2). Il suffit d'appliquer la propriété (1) à l'inclusion $K \subset \overline{K}$.

(2) $\Longrightarrow$ (3). On procède par contraposition. Supposons que L/K ne soit pas séparable. Il existe donc $\alpha \in L \setminus K$ non séparable[3]. Les considérations précédentes appliquées à l'inclusion $K \subset \overline{K}$ montrent alors que le nombre de plongements est $< [L : K(\alpha)][K(\alpha) : K] = [L : K]$, d'où l'implication voulue.

(3) $\Longrightarrow$ (1). Supposons que L/K soit séparable et soit $\iota : K \longrightarrow \overline{K}$ un morphisme d'anneaux. Comme $L \neq K$, on peut choisir un élément $\alpha \in L \setminus K$. Un tel élément est séparable par hypothèse. Ainsi, ι admet $[K(\alpha) : K]$ prolongements à $K(\alpha)$. De plus, l'extension $L/K(\alpha)$ est séparable d'après la remarque 1.10. Par hypothèse de récurrence, il s'ensuit que n'importe quel prolongement de ι à $K(\alpha)$ admet lui-même $[L : K(\alpha)]$ prolongements. Mais alors, ι possède $[K(\alpha) : K][L : K(\alpha)] = [L : K]$ prolongements à L.

Cela achève la récurrence, ainsi que la démonstration. □

1.12. Corollaire. *Soit L/K une extension de degré fini. Alors, on a*

$$|\mathrm{Gal}(L/K)| \leqslant [L : K].$$

3. Les éléments de K sont séparables sur K.

1.13. Remarque. Soit L/K une extension de degré fini. Alors, il existe des éléments $\alpha_1, \ldots, \alpha_n \in L$ tels que $L = K(\alpha_1, \ldots, \alpha_n)$. En effet, on peut par exemple prendre les éléments d'une K-base L. On peut alors déterminer les éléments de $\mathscr{X}(L/K)$ de la manière suivante. Pour tout $i \in [\![0, n]\!]$, posons

$$L_i = K(\alpha_1, \ldots, \alpha_i),$$

si bien que $L_i = L_{i-1}(\alpha_i)$ pour tout $i \in [\![1, n]\!]$.

On trouve tous les prolongements de $K \subset \overline{K}$ à $L_1 = K(\alpha_1)$ en utilisant le théorème XXVI-2.2, puis on étend chaque prolongement $L_1 \longrightarrow \overline{K}$ au corps $L_2 = L_1(\alpha_2)$ en utilisant ce même lemme, et ainsi de suite.

Notation. Soit L/K une extension de degré fini. Écrivons

$$L = K(\alpha_1, \ldots, \alpha_n).$$

Comme on a l'égalité $L = K[\alpha_1, \ldots, \alpha_n]$ (cf. corollaire XXV-3.11), un plongement $\sigma : L \longrightarrow \overline{K}$ est déterminé de manière unique par les images de $\alpha_1, \ldots, \alpha_n$. Si $\beta_i \in \overline{K}$ est l'image de α_i par σ, on notera simplement le morphisme par

$$\sigma : \left| \begin{array}{ccc} \alpha_1 & \rightsquigarrow & \beta_1 \\ & \vdots & \\ \alpha_n & \rightsquigarrow & \beta_n \end{array} \right.$$

1.14. Exemple. Soit $K = \mathbb{Q}$ et soit $L = \mathbb{Q}(\alpha, \beta)$, avec $\alpha = \sqrt[4]{2}$ et $\beta = \sqrt[5]{2}$. D'après le critère d'Eisenstein, on a

$$\mu_{\alpha,K} = X^4 - 2 = (X - \alpha)(X + \alpha)(X - i\alpha)(X + i\alpha).$$

D'après le théorème XXVI-2.2, l'inclusion $K \subset \overline{K}$ possède quatre prolongements

$$\tau_k : \alpha \rightsquigarrow i^k \alpha, \quad k \in [\![0, 3]\!].$$

Passons à l'étape suivante. Pour cela, il faut calculer le polynôme minimal de β sur $K(\alpha)$. D'après l'exemple XXV-3.27, $K(\alpha)/K$ et $K(\beta)/K$ sont linéairement disjointes, et l'on a $[L : K] = 20$. En particulier, on a

$$[L : K(\alpha)] = \frac{[L : K]}{[K(\alpha) : K]} = \frac{20}{4} = 5.$$

Or, $X^5 - 2$ est un polynôme de $K(\alpha)[X]$ unitaire annulant β de degré 5. Comme $\mu_{\beta,K(\alpha)}$ est de degré 5 et divise $X^5 - 2$, on a nécessairement

$$\mu_{\beta,K(\alpha)} = X^5 - 2 = \prod_{m=0}^{m=4} (X - \beta\zeta_5^m).$$

Puisque $\tau_k(\mu_{\beta,K(\alpha)}) = X^5 - 2$, le théorème XXVI-2.2 montre que τ_k admet

cinq prolongements à L, à savoir

$$\sigma_{k,m} : \left| \begin{array}{ccc} \alpha & \rightsquigarrow & i^k\alpha \\ \beta & \rightsquigarrow & \zeta_5^m\beta \end{array} \right. , \quad k \in [\![0,3]\!] , \ m \in [\![0,4]\!] .$$

On a donc exactement vingt plongements de L dans $\overline{K}$. En particulier, L/K est séparable par le théorème 1.11. Notons également que $L \subset \mathbb{R}$. On en déduit alors l'égalité

$$\mathrm{Gal}(L/K) = \{\sigma_{0,0}, \sigma_{2,0}\}.$$

Nous allons maintenant étudier de plus près les extensions séparables.

2. Propriétés des extensions séparables

2.1. Proposition. *Soit L/K une extension. Alors, l'ensemble des éléments de L séparables sur K est une sous-extension de L/K.*

Démonstration. Soit M l'ensemble des éléments de L séparables sur K. Alors, M contient K, donc contient a fortiori 0 et 1. Soient $\alpha, \beta \in M$, et soit $E = K(\alpha, \beta)$. L'extension de corps E/K est de degré fini d'après le corollaire XXV-3.11. Nous allons montrer que E/K est séparable. Cela démontrera en particulier que $-\alpha$, $\alpha + \beta$, $\alpha\beta$ et α^{-1} (lorsque $\alpha \neq 0$) sont séparables, puisque ce sont des éléments de E.

Pour cela, comptons le nombre de plongements de E/K. Puisque α est séparable sur K, l'inclusion $K \subset \overline{K}$ admet $[K(\alpha) : K]$ prolongements au corps $K(\alpha)$ d'après le lemme 1.6. Puisque β est séparable sur K, il est aussi séparable sur $K(\alpha)$ par la remarque 1.10. Chaque plongement de $K(\alpha)$ admet alors $[E : K(\alpha)]$ prolongements à E, toujours d'après le lemme 1.6. Ainsi, E/K admet exactement

$$[K(\alpha) : K] \times [E : K(\alpha)] = [E : K]$$

plongements. Par le théorème 1.11, E/K est séparable, d'où le résultat. □

On en déduit alors le résultat suivant.

2.2. Corollaire. *Soit L/K une extension, et soit $\mathscr{S}$ une partie de L. Alors, l'extension $K(\mathscr{S})/K$ est séparable si, et seulement si, tout élément de $\mathscr{S}$ est séparable sur K.*

Démonstration. Soit M l'ensemble des éléments de L qui sont séparables sur K. Supposons que tout élément de $\mathscr{S}$ soit séparable sur K. Alors, $\mathscr{S}$ est contenue dans M. Comme M est un sous-corps de L contenant K d'après la proposition précédente, on a $K(\mathscr{S}) \subset M$. Autrement dit, $K(\mathscr{S})/K$ est séparable. L'autre implication est immédiate. □

2.3. Corollaire. *Soit L/K une extension, et soient $E/K, E'/K$ deux sous-extensions de L/K. Si E/K est séparable, alors EE'/E' est séparable.*

Démonstration. Puisque tout élément de E est séparable sur K, tout élément de E est aussi séparable sur E' par la remarque 1.10. Ainsi, tout élément de E est séparable sur E'. Par le corollaire précédent, tout élément de la sous-extension M/E' de L/E' engendrée par E est séparable. Or, cette extension n'est rien d'autre que EE', comme on l'a déjà constaté. Cela achève la démonstration. □

2.4. Corollaire. *Soit L/K une extension, et soient $E/K, E'/K$ deux sous-extensions de L/K. Si E/K et E'/K sont séparables, alors EE'/K est séparable.*

Démonstration. Il suffit d'appliquer le corollaire 2.2 à $\mathscr{S} = E \cup E'$. □

On a enfin le résultat suivant.

2.5. Proposition. *Soit $K \subset M \subset L$ une tour d'extensions. Alors, L/K est séparable si, et seulement si, M/K et L/M sont séparables.*

Démonstration. Le sens direct est donné par la remarque 1.10. Supposons maintenant que M/K et L/M soient séparables. En particulier, les extensions M/K et L/K sont algébriques par la proposition XXV-3.19. Soit $\alpha \in L$. Par hypothèse, α est séparable sur M. Écrivons

$$\mu_{\alpha,M} = X^n + a_{n-1}X^{n-1} + \cdots + a_0 \in M[X],$$

et posons $E = K(a_0, \ldots, a_{n-1})$. Cette extension est de degré fini par le corollaire XXV-3.11. De même, $E(\alpha)/K$ est de degré fini. Nous allons montrer que cette extension est séparable.

Puisque $\mu_{\alpha,M} \in E[X]$ est irréductible dans $M[X]$, il est aussi irréductible dans $E[X]$, et l'on a donc $\mu_{\alpha,E} = \mu_{\alpha,M}$. Ainsi, α est séparable sur E. Puisque les éléments $a_0, \ldots, a_{n-1} \in M$ sont séparables sur K, E/K est séparable sur K d'après le corollaire 2.2, et de degré fini. Le nombre de plongements de E/K est donc égal à $[E : K]$ d'après le théorème 1.11. Comme α est séparable sur E, chaque plongement de E admet $[E(\alpha) : E]$ extensions à $E(\alpha)$ par le lemme 1.6. Ainsi, $E(\alpha)/K$ admet $[E(\alpha) : K]$ plongements, et $E(\alpha)/K$ est donc séparable. En particulier, α est séparable sur K, ce qui achève la démonstration. □

Nous continuons maintenant en démontrant qu'une extension séparable de degré fini est monogène. On commence par un lemme.

2.6. Lemme. *Soit L/K une extension de corps séparable de degré $n \geqslant 1$, et soient $\sigma_1, \ldots, \sigma_n$ les n plongements de L/K. Alors, pour tout $\alpha \in L$, on a $L = K(\alpha)$ si, et seulement si, les n éléments $\sigma_1(\alpha), \ldots, \sigma_n(\alpha) \in \overline{K}$ sont distincts.*

Démonstration. Remarquons d'abord que la restriction de chaque plongement de L/K à $K(\alpha)$ fournit un plongement de $K(\alpha)/K$. Ainsi, si les éléments $\sigma_1(\alpha), \ldots, \sigma_n(\alpha) \in \overline{K}$ sont tous distincts, $K(\alpha)/K$ possède au moins n plongements distincts. D'après le lemme 1.6 par exemple, on a donc $n \leqslant [K(\alpha) : K]$.

Puisque l'on a aussi $[K(\alpha) : K] \leqslant [L : K] = n$, on en déduit

$$[K(\alpha) : K] = n = [L : K].$$

Ainsi, $L = K(\alpha)$. Si maintenant $L = K(\alpha)$, puisque chaque plongement est uniquement déterminé par l'image de α (cf. théorème XXVI-2.2) et que les plongements $\sigma_1, \ldots, \sigma_n$ sont deux à deux distincts par hypothèse, on en déduit que les n éléments $\sigma_1(\alpha), \ldots, \sigma_n(\alpha) \in \overline{K}$ sont distincts, d'où le résultat souhaité. □

Énonçons à présent le théorème de l'élément primitif.

2.7. Théorème (de l'élément primitif). *Soit L/K une extension séparable de degré fini. Alors, il existe $\alpha \in L$ tel que $L = K(\alpha)$.*

Démonstration. Si K est fini, on applique le lemme XXVI-2.9.

Supposons maintenant que K soit infini. L'extension L/K admet donc n plongements $\sigma_1, \ldots, \sigma_n$ deux à deux distincts, où $n = [L : K]$. Puisque L est un K-espace vectoriel de dimension finie n, il est engendré par un nombre fini d'éléments. On peut donc écrire

$$L = K(\alpha_1, \ldots, \alpha_m),$$

pour certains $\alpha_1, \ldots, \alpha_m \in L$. Si $m = 1$, il n'y a rien à faire. Si $m \geqslant 2$, une récurrence facile montre qu'il suffit de considérer le cas $m = 2$. Supposons donc que $L = K(\alpha_1, \alpha_2)$, et posons

$$P = \prod_{1 \leqslant i < j \leqslant n} \big(\sigma_i(\alpha_1) - \sigma_j(\alpha_1) + \big(\sigma_i(\alpha_2) - \sigma_j(\alpha_2)\big)X\big) \in \overline{K}[X].$$

Montrons que l'on a $P \neq 0$. Si $P = 0$, cela signifie qu'un des facteurs de degré 1 est nul, et donc qu'il existe $1 \leqslant i < j \leqslant n$ tels que

$$\sigma_i(\alpha_1) = \sigma_j(\alpha_1) \quad \text{et} \quad \sigma_i(\alpha_2) = \sigma_j(\alpha_2).$$

Mais, comme α_1 et α_2 engendrent L comme anneau, on en déduit $\sigma_i = \sigma_j$, d'où une contradiction.

Ainsi, $P \neq 0$, et puisque K est infini, il existe $\lambda \in K$ tel que $P(\lambda) \neq 0$. Pour tout $1 \leqslant i < j \leqslant n$, on a donc $\sigma_i(\alpha_1) - \sigma_j(\alpha_1) + (\sigma_i(\alpha_2) - \sigma_j(\alpha_2))\lambda \neq 0$, c'est-à-dire $\sigma_i(\alpha_1 + \lambda\alpha_2) \neq \sigma_j(\alpha_1 + \lambda\alpha_2)$.

Posons $\alpha = \alpha_1 + \lambda\alpha_2$. Les n éléments $\sigma_1(\alpha), \ldots, \sigma_n(\alpha)$ sont distincts par construction, et donc $L = K(\alpha)$, d'après le lemme 2.6. □

2.8. Définition. Soit L/K une extension monogène. Un élément $\alpha \in L$ est dit *primitif* si $L = K(\alpha)$.

2.9. Remarques

(1) La réciproque de ce théorème est bien entendu fausse, comme le montre déjà l'exemple 1.9 (3).

(2) Un examen attentif de la démonstration montre que l'on a le résultat plus précis suivant : soit L/K une extension séparable de degré fini, et soient $\alpha_1, \ldots, \alpha_m \in L$ tels que $L = K(\alpha_1, \ldots, \alpha_m)$. Si K est infini, il existe un élément primitif $\alpha \in L$ de la forme

$$\alpha = \lambda_1\alpha_1 + \cdots + \lambda_n\alpha_n, \quad \lambda_i \in K.$$

2.10. Exemple. Soit $K = \mathbb{Q}$ et soit $L = \mathbb{Q}(\alpha, \beta)$, avec $\alpha = \sqrt[4]{2}$ et $\beta = \sqrt[5]{2}$. D'après l'exemple 1.14, on sait que L/K est séparable, et que les plongements sont donnés par

$$\sigma_{k,m} : \left| \begin{array}{ccc} \alpha & \rightsquigarrow & i^k\alpha \\ \beta & \rightsquigarrow & \zeta_5^m\beta \end{array} \right. , \quad k \in [\![0,3]\!],\ m \in [\![0,4]\!].$$

On vérifie alors, en utilisant le lemme 2.6, que $L = K(\alpha + \beta)$.

On finit ce paragraphe en faisant le lien entre plongements et polynômes minimaux.

2.11. Lemme. *Soit L/K une extension séparable de degré fini. Alors, pour tout $\alpha \in L$, on a*

$$\prod_{\sigma \in \mathscr{X}(L/K)} (X - \sigma(\alpha)) = \mu_{\alpha,K}^{[L:K(\alpha)]}.$$

En particulier, si $\alpha \in L$ est un élément primitif, on a

$$\prod_{\sigma \in \mathscr{X}(L/K)} (X - \sigma(\alpha)) = \mu_{\alpha,K}.$$

Démonstration. Posons $M = K(\alpha)$, et soit $n = [M : K]$. Par hypothèse, α est séparable sur K. Soient $\alpha_1, \ldots, \alpha_n \in \overline{K}$ les racines de $\mu_{\alpha,K}$ dans $\overline{K}$.

On peut donc écrire

$$\mu_{\alpha,K} = \prod_{i=1}^{n}(X - \alpha_i) \in \overline{K}[X].$$

Le théorème XXVI-2.2 montre alors que les plongements sont les morphismes

$$\sigma_i : \alpha \rightsquigarrow \alpha_i, \quad i \in [\![1, n]\!].$$

Comme L/K est séparable, L/M est séparable (cf. remarque 1.10) et chaque σ_i admet exactement $[L : M]$ prolongements à L.

Autrement dit, pour un $i \in [\![1, n]\!]$ fixé, il y a exactement $[L : M]$ plongements dont la restriction à M est σ_i. Pour un tel plongement σ, on a alors

$$\sigma(\alpha) = \sigma_i(\alpha) = \alpha_i,$$

puisque $\alpha \in M$. Ainsi, on obtient

$$\prod_{\sigma \in \mathscr{X}(L/K)} \big(X - \sigma(\alpha)\big) = \Big(\prod_{i=1}^{n}(X - \alpha_i)\Big)^{[L:K(\alpha)]} = \mu_{\alpha,K}^{[L:K(\alpha)]},$$

d'où le résultat. □

2.12. Corollaire. *Soit L/K une extension séparable de degré fini. Alors, pour tout $\alpha \in L$, on a*

$$\sigma(\alpha) = \alpha, \quad \textit{pour tout } \sigma \in \mathscr{X}(L/K) \quad \Longleftrightarrow \quad \alpha \in K.$$

Démonstration. Si $\alpha \in L$ est fixé par tout élément de $\mathscr{X}(L/K)$, le lemme précédent montre que

$$\mu_{\alpha,K}^{[L:K(\alpha)]} = (X - \alpha)^{[L:K]}.$$

En particulier, $\mu_{\alpha,K}$ admet une unique racine dans $\overline{K}$, à savoir α. Comme le polynôme $\mu_{\alpha,K}$ est séparable, on en déduit que

$$\mu_{\alpha,K} = X - \alpha \in \overline{K}[X],$$

et puisque $\mu_{\alpha,K}$ est à coefficients dans K, on obtient $\alpha \in K$. Le sens indirect est clair, puisqu'un plongement est K-linéaire et envoie 1 sur 1. Cela achève la démonstration. □

2.13. Remarque. Soit L/K une extension de degré fini (pas nécessairement séparable), et soit $\alpha \in L$ un élément séparable. Le lemme 2.11 montre en particulier que l'on a

$$\prod_{\tau \in \mathscr{X}(K(\alpha)/K)} \big(X - \tau(\alpha)\big) = \mu_{\alpha,K}.$$

On peut alors en déduire une méthode pour calculer le polynôme minimal de α, dès lors que l'on connaît tous les plongements de L dans $\overline{K}$.

On calcule les éléments $\sigma(\alpha), \sigma \in \mathscr{X}(L/K)$, et l'on choisit une famille

$$\sigma_1(\alpha), \ldots, \sigma_r(\alpha)$$

d'éléments deux à deux distincts de cardinal maximal. Alors, on a

$$\mu_{\alpha,K} = \prod_{i=1}^{r} \big(X - \sigma_i(\alpha)\big).$$

En effet, les éléments $\tau(\alpha), \tau \in \mathscr{X}(K(\alpha)/K)$, sont tous de la forme $\sigma(\alpha)$, avec $\sigma \in \mathscr{X}(L/K)$, puisque chaque plongement de $K(\alpha)$ se prolonge[4] à L. La maximalité de r assure alors l'égalité

$$\{\sigma_1(\alpha), \ldots, \sigma_r(\alpha)\} = \{\tau(\alpha) \mid \tau \in \mathscr{X}\big(K(\alpha)/K\big)\},$$

d'où le résultat.

2.14. Exemple. Soit $L = \mathbb{Q}(\sqrt{2}, \sqrt{3})$. On a

$$\mu_{\sqrt{2},\mathbb{Q}} = X^2 - 2 = (X - \sqrt{2})(X + \sqrt{2}),$$

et l'inclusion $\mathbb{Q} \subset \overline{\mathbb{Q}}$ se prolonge en deux plongements, à savoir

$$\tau_m : \sqrt{2} \rightsquigarrow (-1)^k \sqrt{2}, \quad k = 0, 1.$$

Montrons que $\sqrt{3} \notin \mathbb{Q}(\sqrt{2})$. Supposons que $\sqrt{3} = a + b\sqrt{2}$, $a, b \in \mathbb{Q}$. En élevant au carré, on obtient

$$3 = a^2 + 2b^2 + 2ab\sqrt{2}.$$

On en déduit $ab = 0$. On a donc $a = 0$, auquel cas $3 = 2b^2$, ou $b = 0$, auquel cas $3 = a^2$. Dans les deux cas, on obtient une contradiction en comparant les valuations 3-adiques. Mais alors, $X^2 - 3 \in \mathbb{Q}[X] \subset \mathbb{Q}(\sqrt{2})[X]$ est unitaire, irréductible, et annule $\sqrt{3}$. On a donc $\mu_{\sqrt{3},\mathbb{Q}(\sqrt{2})} = X^2 - 3$. Comme $\tau_k(X^2 - 3) = X^2 - 3$ pour $k = 0, 1$, les plongements de M dans $\overline{\mathbb{Q}}$ sont

$$\sigma_{k,m} : \left| \begin{array}{lll} \sqrt{2} & \rightsquigarrow & (-1)^k \sqrt{2} \\ \sqrt{3} & \rightsquigarrow & (-1)^m \sqrt{3} \end{array} \right. , \quad \text{pour tous } k, m \in \{0, 1\}.$$

Soit $\alpha = \sqrt{2} + \sqrt{3}$. Les images de α par les divers plongements sont distinctes, et valent

$$\sqrt{2} + \sqrt{3}, \quad \sqrt{2} - \sqrt{3}, \quad -\sqrt{2} + \sqrt{3}, \quad -\sqrt{2} - \sqrt{3}.$$

On a alors

$$\mu_{\alpha,\mathbb{Q}} = (X - \sqrt{2} - \sqrt{3})(X + \sqrt{2} - \sqrt{3})(X - \sqrt{2} + \sqrt{3})(X + \sqrt{2} + \sqrt{3}),$$

soit $\mu_{\alpha,\mathbb{Q}} = X^4 - 10X^2 + 1$.

4. On peut remarquer que $L/K(\alpha)$ est de degré fini, et procéder comme dans la remarque 1.13 pour prolonger τ, ou utiliser la proposition XXVI-3.5.

On en déduit a posteriori que ce polynôme-là est irréductible, ce que l'on savait déjà par la proposition XXI-4.4.

3. Polynômes séparables

Nous allons maintenant étudier à quelques conditions un polynôme de $K[X]$ est séparable, en mettant l'accent sur les polynômes irréductibles.

Commençons par un lemme.

3.1. Lemme. *Soient $f, g \in K[X]$ et soit $\iota : K \longrightarrow L$ un morphisme d'anneaux. Alors, on a*

$$\operatorname{pgcd}\big(\iota(f), \iota(g)\big) = \iota\big(\operatorname{pgcd}(f, g)\big).$$

Démonstration. Si P_1, $P_2 \in K[X]$, avec $P_2 \neq 0$, soit $P_1 = P_2Q + R$ la division euclidienne de P_1 par P_2. On a donc $\deg(R) < \deg(P_2)$. Le morphisme d'anneaux ι étant injectif, pour tout $S \in K[X]$, les polynômes S et $\iota(S)$ ont même degré. Comme on a

$$\iota(P_1) = \iota(P_2Q + R) = \iota(P_2)\iota(Q) + \iota(R),$$

cela montre que le polynôme $\iota(R) \in L[X]$ est le reste de la division euclidienne de $\iota(P)$ par $\iota(Q)$. De plus, $\iota(R)$ est nul si, et seulement si, R l'est. En particulier, si $R_0, R_1, \ldots, R_n$ sont les restes non nuls obtenus dans l'algorithme d'Euclide associé à f et g, $\iota(R_0), \iota(R_1), \ldots, \iota(R_n)$ sont les restes non nuls obtenus dans l'algorithme d'Euclide associé à $\iota(f)$ et $\iota(g)$. Par conséquent, $\iota(R_n)$ est un pgcd de $\iota(f)$ et $\iota(g)$, ce qui démontre le résultat voulu en normalisant les pgcd. □

On continue par une caractérisation de la séparabilité en termes de polynôme dérivé, bien connue dans le cas des polynômes à coefficients réels.

3.2. Lemme. *Soit K un corps, et soit $P \in K[X]$ un polynôme non constant. Alors, P est séparable si, et seulement si, P et P' sont premiers entre eux.*

Démonstration. On peut supposer P unitaire sans perte de généralité. Le lemme précédent, appliqué à l'inclusion $K \subset \overline{K}$, montre que les pgcd unitaires de P et P' calculés dans $K[X]$ et $\overline{K}[X]$, sont égaux. Quitte à remplacer K par $\overline{K}$, on peut donc aussi supposer que P est scindé sur K. Écrivons

$$P = \prod_{i=1}^{r} (X - \alpha_i)^{m_i}, \quad \alpha_i \in K,\ m_i \geqslant 1,\ r \geqslant 1.$$

On a alors

$$P' = \sum_{i=1}^{r} m_i (X - \alpha_i)^{m_i - 1} \prod_{j \neq i} (X - \alpha_j)^{m_j}.$$

Un diviseur irréductible unitaire éventuel de $\operatorname{pgcd}(P, P')$ est un diviseur irréductible unitaire de P. C'est donc un des polynômes $X - \alpha_k$, $k \in [\![1, r]\!]$.

Soit $k \in [\![1, r]\!]$. Remarquons que pour tout $i \neq k$, on a

$$X - \alpha_k \mid \prod_{j \neq i} (X - \alpha_j)^{m_j}.$$

Ainsi, on a

$$\begin{aligned} X - \alpha_k \mid P' &\iff X - \alpha_k \mid m_k (X - \alpha_k)^{m_k - 1} \prod_{j \neq k} (X - \alpha_j)^{m_j} \\ &\iff X - \alpha_k \mid m_k (X - \alpha_k)^{m_k - 1}, \end{aligned}$$

la dernière équivalence provenant du lemme de Gauss.

Si P est séparable, on a $m_k = 1$ pour tout $k \in [\![1, r]\!]$, et cette dernière condition de divisibilité n'a jamais lieu. Ainsi, P et P' sont premiers entre eux dans ce cas.

Si P n'est pas séparable, il existe $k \in [\![1, r]\!]$ tel que $m_k > 1$. La condition précédente montre alors que $X - \alpha_k \mid P'$. Dans ce cas, P et P' ne sont pas premiers entre eux, d'où le résultat. □

3.3. Remarque. Soit $\iota : K \longrightarrow L$ un morphisme d'anneaux. Les deux résultats précédents permettent de démontrer que $P \in K[X]$ est séparable si, et seulement si, $\iota(P) \in L[X]$ est séparable, et cela sans utiliser la notion de discriminant d'un polynôme.

En effet, en remarquant que $\iota(P') = \iota(P)'$, on obtient alors

$$\begin{aligned} \operatorname{pgcd}(P, P') = 1 &\iff \iota\big(\operatorname{pgcd}(P, P')\big) = \iota(1) \\ &\iff \iota\big(\operatorname{pgcd}(P, P')\big) = 1 \\ &\iff \operatorname{pgcd}\big(\iota(P), \iota(P')\big) = 1 \\ &\iff \operatorname{pgcd}\big(\iota(P), \iota(P)'\big) = 1, \end{aligned}$$

d'où le résultat.

3.4. Corollaire. *Soit K un corps, et soit $P \in K[X]$ un polynôme irréductible. Alors, P n'est pas séparable si, et seulement si, $P' = 0$.*

Démonstration. On peut toujours supposer P unitaire. Puisque P est irréductible, les diviseurs unitaires de P sont 1 ou P. On a donc

$$\operatorname{pgcd}(P, P') = 1 \quad \text{ou} \quad P.$$

D'après le lemme précédent, on a

$$P \text{ n'est pas séparable} \iff \operatorname{pgcd}(P, P') = P \iff P \mid P'.$$

Mais, $\deg(P') < \deg(P)$. La condition $P \mid P'$ est alors réalisée si, et seulement si, $P' = 0$. Cela achève la démonstration. □

On peut maintenant décortiquer la structure des polynômes irréductibles séparables.

3.5. Lemme. *Soit K un corps.*

(1) *Si* $\mathrm{car}(K) = 0$, *tout polynôme irréductible de* $K[X]$ *est séparable.*

(2) *Si* $\mathrm{car}(K) = p > 0$, *et si* $P \in K[X]$ *est irréductible, alors* P *n'est pas séparable si, et seulement si, il existe* $Q \in K[X]$ *tel que* $P = Q(X^p)$.

Démonstration. Soit $P \in K[X]$ un polynôme irréductible de degré $n \geqslant 1$. Si $P = a_n X^n + \cdots + a_1 X + a_0$, avec $a_n \in K^\times$, on a

$$P' = n a_n X^{n-1} + \cdots + 2a_2 X + a_1.$$

Ainsi, on a

$$P' = 0 \iff ka_k = 0, \text{ pour tout } k \in [\![1, n]\!].$$

Si $\mathrm{car}(K) = 0$, on a $na_n \neq 0$ et $P' \neq 0$. Par conséquent, P est séparable par le corollaire 3.4.

Supposons que $\mathrm{car}(K) = p > 0$. Si $p \mid k$, alors $k = pm, m \in \mathbb{Z}$ et l'on a

$$ka_k = p(ma_k) = 0,$$

car K est de caractéristique p. Si $p \nmid k$, alors $k \cdot 1$ est inversible dans K (puisque le sous-anneau engendré par 1 est isomorphe à $\mathbb{Z}/p\mathbb{Z}$), et dans ce cas, on a[5]

$$ka_k = 0 \iff a_k = 0.$$

Ainsi, $P' = 0$ si, et seulement si, il existe $Q \in K[X]$ tel que $P = Q(X^p)$, ce qui démontre le résultat voulu. □

En guise d'application, nous caractérisons les corps dont toute extension algébrique est séparable. De tels corps méritent un nom bien à eux.

3.6. Définition. Un corps K est dit *parfait* si toute extension algébrique de K est séparable.

On a alors le théorème suivant.

3.7. Théorème. *Soit K un corps. Alors, K est parfait si, et seulement si, une des deux conditions suivantes est réalisée :*

(1) $\mathrm{car}(K) = 0$;

(2) $\mathrm{car}(K) = p > 0$ *et* $K = K^p$, *où* $K^p = \{x^p \mid x \in K\}$.

En particulier, tout corps fini est parfait.

5. On peut aussi remarquer que k et p sont premiers entre eux. Si $u, v \in \mathbb{Z}$ sont deux entiers tels que $uk + vp = 1$, on a alors

$$a_k = u(ka_k) + v(pa_k) = u \cdot 0 + v \cdot 0 = 0.$$

Démonstration. Remarquons qu'un corps K est parfait si, et seulement si, tout élément algébrique sur K est séparable, ou encore si, et seulement si, tout polynôme irréductible unitaire de $K[X]$ est séparable.

Si $\mathrm{car}(K) = 0$, tout polynôme irréductible unitaire de $K[X]$ est séparable par le lemme 3.5 (1).

On se place maintenant dans le cas où $\mathrm{car}(K) = p > 0$. Supposons tout d'abord que $K = K^p$, et soit P un polynôme irréductible unitaire de $K[X]$. Si P n'est pas séparable, il existe $Q \in K[X]$ tel que $P = Q(X^p)$, par le lemme 3.5 (2). On peut donc écrire

$$P = a_n X^{pn} + \cdots + a_1 X^p + a_0, \quad a_k \in K,\ a_n \in K^\times.$$

Par hypothèse, pour tout $k \in [\![0, n]\!]$, il existe $b_k \in K$ tel que $a_k = b_k^p$. Le lemme XV-1.18 montre alors que l'on a

$$P = (b_n X^n + \cdots + b_1 X + b_0)^p.$$

En particulier, P n'est pas irréductible, d'où une contradiction. Ainsi, P est séparable.

Supposons maintenant que $K \neq K^p$. Il existe donc $a \in K \setminus K^p$. Montrons que $P = X^p - a$ est irréductible, non séparable. Soit $\alpha \in \overline{K}$ une racine de P. On a alors $\alpha^p = a$, et ainsi

$$P = X^p - \alpha^p = (X - \alpha)^p.$$

En particulier, P n'est pas séparable. Supposons que P ne soit pas irréductible, et soit $Q \in K[X]$ un diviseur unitaire de degré $d \in [\![1, p-1]\!]$. Alors, on a $Q = (X - \alpha)^d \in \overline{K}[X]$. Comme Q est à coefficients dans K, le coefficient de X^{d-1} est dans K. Autrement dit, $-d\alpha \in K$. Comme $d \in [\![1, p-1]\!]$, l'élément $d \cdot 1_K$ est inversible dans K, on en déduit que $-\alpha \in K$. Par conséquent, $\alpha \in K$. Mais alors, $a = \alpha^p \in K^p$, d'où une contradiction. Ainsi, P est irréductible et non séparable.

Cela démontre la première partie du théorème.

Supposons maintenant que K soit un corps à p^d éléments, où p est premier et $d \geqslant 1$. D'après le lemme XXI-6.3, pour tout $x \in K$, on a

$$x = x^{p^d} = \left(x^{p^{d-1}}\right)^p \in K^p.$$

Ainsi, $K = K^p$. Comme $p = \mathrm{car}(K)$, on en déduit que K est parfait d'après la première partie. Cela achève la démonstration. □

3.8. Remarque. Le théorème précédent montre deux choses :

(i) la séparabilité d'une extension L/K est automatique lorsque K est de caractéristique nulle ou un corps fini ;

(ii) il faut vraiment se fatiguer pour construire des éléments algébriques sur K non séparables. En fait, si $\overline{K}$ contient un élément non séparable, alors K est une extension transcendante de $\mathbb{F}_p$, où p est un nombre premier. En effet, si le corps K n'est pas parfait, on a $\mathrm{car}(K) = p > 0$. Dans ce cas, K est une extension de $\mathbb{F}_p$. Supposons que $K/\mathbb{F}_p$ soit algébrique. Si L/K est une extension algébrique, alors $L/\mathbb{F}_p$ est une extension algébrique par la proposition XXV-3.19, donc séparable puisque $\mathbb{F}_p$ est parfait. Mais alors, L/K est aussi séparable par la remarque 1.10. Ainsi, si l'extension $K/\mathbb{F}_p$ est algébrique, K est parfait. Cela montre le fait annoncé.

4. Traces et normes

Pour finir ce chapitre, nous allons introduire la trace et la norme d'une extension, qui généralisent respectivement la partie réelle et le module d'un nombre complexe [6], et qui s'avèrent extrêmement utiles pour les calculs. Les définitions et démonstrations nécessitent quelques notations de base d'algèbre linéaire (définition et propriétés du polynôme caractéristique, trace, déterminant). Le lecteur désireux de se rafraîchir la mémoire pourra consulter avec profit le paragraphe XXX-1.

Comme ce n'est pas plus compliqué, nous allons les définir pour une algèbre associative unitaire de dimension finie.

4.1. Définition. Soit E une K-algèbre associative unitaire de dimension finie. Pour tout $x \in E$, on appelle *trace* et *norme* de $x \in E$, la trace et le déterminant de l'application K-linéaire

$$\begin{aligned} \ell_x : E &\longrightarrow E \\ y &\longmapsto xy. \end{aligned}$$

On les note respectivement $\mathrm{Tr}_{E/K}(x)$ et $N_{E/K}(x)$.

4.2. Exemples

(1) Soit $E = K^n$. Si $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, la matrice représentative de la multiplication à gauche par $x \in E$ dans la base canonique est la matrice diagonale

$$\begin{pmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{pmatrix}.$$

6. Ou presque...

On a alors

$$\mathrm{Tr}_{E/K}(x) = x_1 + \cdots + x_n \quad \text{et} \quad N_{E/K}(x) = x_1 \ldots x_n.$$

(2) Soit $E = \mathbb{Q}(\sqrt{d})$, où $d \in K$ n'est pas un carré dans K. Dans ce cas, la famille $(1, \sqrt{d})$ est une K-base de E. Si $x = a + b\sqrt{d}$, $a, b \in K$, la matrice représentative de la multiplication à gauche par $x \in E$ dans cette base est

$$\begin{pmatrix} a & bd \\ b & a \end{pmatrix}.$$

Ainsi, on a

$$\mathrm{Tr}_{K(\sqrt{d})/K}(a + b\sqrt{d}) = 2a \quad \text{et} \quad N_{K(\sqrt{d})/K}(a + b\sqrt{d}) = a^2 - db^2.$$

En particulier, si $K = \mathbb{R}$ et $d = -1$, on a $E = \mathbb{C}$, et l'on obtient

$$\mathrm{Tr}_{\mathbb{C}/\mathbb{R}}(z) = 2\,\mathrm{Re}(z) \quad \text{et} \quad N_{\mathbb{C}/\mathbb{R}}(z) = |z|^2.$$

Au passage, le lecteur notera qu'il aura utilisé la norme de $\mathbb{Q}(\sqrt{d})/\mathbb{Q}$ sans le savoir, pour résoudre diverses questions d'arithmétique dans les anneaux du type $\mathbb{Z}[\sqrt{d}]$ ou $\mathbb{Z}\left[\frac{-1+\sqrt{d}}{2}\right]$ (cf. chapitre XVIII et ses exercices).

Le lemme suivant donne quelques propriétés de la trace et de la norme.

4.3. Lemme. *Soit E une K-algèbre associative unitaire, de dimension finie. Alors :*

(1) *l'application $\mathrm{Tr}_{E/K} : E \longrightarrow K$ est une forme linéaire ;*

(2) *pour tous $x, y \in E$, on a*

$$\mathrm{Tr}_{E/K}(xy) = \mathrm{Tr}_{E/K}(yx) \quad \text{et} \quad N_{E/K}(xy) = N_{E/K}(x)N_{E/K}(y)$$

(3) *pour tout $\lambda \in K$, on a $\mathrm{Tr}_{E/K}(\lambda) = n \cdot \lambda$ et $N_{E/K}(\lambda) = \lambda^n$, où n est la dimension de E sur K ;*

(4) *pour tout $x \in E$, on a $N_{E/K}(x) \neq 0$ si, et seulement si, $x \in E^\times$.*

Démonstration. Une application des définitions montre que

$$\begin{aligned} \mu_{E/K} \colon E &\longrightarrow \mathscr{L}(E) \\ x &\longmapsto \ell_x \end{aligned}$$

est K-linéaire. Or, la trace $\mathrm{tr} : \mathscr{L}(E) \longrightarrow K$ est également K-linéaire. Ainsi, $\mathrm{Tr}_{E/K} = \mathrm{tr} \circ \mu_{E/K}$ est aussi K-linéaire, d'où (1).

De plus, pour tous $x, y \in E$, on a $\ell_{xy} = \ell_x \circ \ell_y$. Les propriétés usuelles de la trace et du déterminant d'un endomorphisme donnent alors (2). Si maintenant $\lambda \in K$, on a $\ell_\lambda = \lambda \mathrm{Id}_E$, et (3) s'en déduit immédiatement.

Montrons (4). Si $x \in E^\times$, alors ℓ_x est inversible, d'inverse $\ell_{x^{-1}}$. En particulier,

$$N_{E/K}(x) = \det(\ell_x) \neq 0.$$

Supposons maintenant que $N_{E/K}(x) \neq 0$. Alors, ℓ_x est inversible, donc surjective. Il existe alors $y \in E$ tel que $\ell_x(y) = xy = 1_E$. Mais alors, on a

$$x(yx) = (xy)x = x = x \cdot 1_E.$$

Comme ℓ_x est bijective, on en déduit $yx = 1_E$. Ainsi, $xy = yx = 1_E$, ce qui démontre que $x \in E^\times$. Cela achève la démonstration. □

Nous allons maintenant nous intéresser de plus près aux propriétés de la trace et de la norme des extensions de degré fini. On commence par un lemme qui permet de calculer la trace et la norme d'un élément pourvu que l'on connaisse son polynôme minimal.

4.4. Lemme. *Soit L/K une extension de degré fini. Alors, pour tout $x \in L$, on a*

$$\chi_{\ell_x} = \mu_{x,K}^{[L:K(x)]}.$$

En particulier :

(1) *si $\mu_{x,K} = X^n + a_{n-1}X^{n-1} + \cdots + a_0$, alors $\mathrm{Tr}_{K(x)/K}(x) = -a_{n-1}$ et $N_{K(x)/K}(x) = (-1)^n a_0$;*

(2) $\mathrm{Tr}_{L/K}(x) = [L:K(x)] \cdot \mathrm{Tr}_{K(x)/K}(x)$ *et* $N_{L/K}(x) = \big(N_{K(x)/K}(x)\big)^{[L:K(x)]}$.

Démonstration. Soit $n = [K(x) : K]$ et soit $m = [L : K(x)]$.

Soit $\mathbf{f} = (f_1, \ldots, f_m)$ une $K(x)$-base de L. Comme $\mathbf{e} = (1, x, \ldots, x^{n-1})$ est une K-base de $K(x)$, la famille

$$\mathbf{g} = (f_1, xf_1, \ldots, x^{n-1}f_1, \ldots, f_m, xf_m, \ldots, x^{n-1}f_m)$$

est une K-base[7] de L.

On vérifie alors aisément l'égalité

$$\mathrm{Mat}(\ell_x ; \mathbf{g}) = \begin{pmatrix} \mathrm{Mat}(\ell_x; \mathbf{e}) & & \\ & \ddots & \\ & & \mathrm{Mat}(\ell_x; \mathbf{e}) \end{pmatrix},$$

où ℓ_x est vu dans le membre de droite comme un endomorphisme de $K(x)$.

7. C'est un cas particulier d'un fait général établi dans la démonstration de la formule de multiplicativité des degrés.

Or, on vérifie aisément que $\mathrm{Mat}(\ell_x\,;\mathbf{e})$ n'est rien d'autre que la matrice compagnon du polynôme $\mu_{x,K}$. Le polynôme caractéristique de $\ell_x \in \mathscr{L}\big(K(x)\big)$ est donc $\mu_{x,K}$ d'après l'exemple XXX-1.13. On en déduit alors l'égalité voulue.

Remarquons au passage que l'on a obtenu que le polynôme caractéristique de $\ell_x \in \mathscr{L}\big(K(x)\big)$ est $\mu_{x,K}$. Le lemme XXX-1.11 donne alors le point (1). Il suffit ensuite d'appliquer successivement ce même lemme et le point (1) à l'égalité $\chi_{\ell_x} = \mu_{x,K}^{[L:K(x)]}$ pour obtenir

$$\mathrm{Tr}_{L/K}(x) = -[L : K(x)]\cdot a_{n-1} = [L : K(x)]\cdot \mathrm{Tr}_{K(x)/K}(x),$$

ainsi que

$$N_{L/K}(x) = (-1)^{n[L:K(x)]} a_0^{[L:K(x)]} = \big(N_{K(x)/K}(x)\big)^{[L:K(x)]}.$$

Cela achève la démonstration. □

La proposition suivante donne un autre moyen de calculer la trace et la norme dans une extension séparable.

4.5. Proposition. *Soit L/K une extension séparable de degré fini. Alors, pour tout $x \in L$, on a*

$$\mathrm{Tr}_{L/K}(x) = \sum_{\sigma\in\mathscr{X}(L/K)} \sigma(x) \quad \textit{et} \quad N_{L/K}(x) = \prod_{\sigma\in\mathscr{X}(L/K)} \sigma(x)$$

Démonstration. Le lemme précédent donne $\chi_{\ell_x} = \mu_{x,K}^{[L:K(x)]}$ pour tout $x \in L$. Le lemme 2.11 nous fournit alors l'égalité

$$\chi_{\ell_x} = \prod_{\sigma\in\mathscr{X}(L/K)} \big(X - \sigma(x)\big).$$

D'après le lemme XXX-1.15, on obtient alors

$$\mathrm{Tr}_{L/K}(x) = \mathrm{tr}(\ell_x) = \sum_{\sigma\in\mathscr{X}(L/K)} \sigma(x),$$

ainsi que

$$N_{L/K}(x) = \det(\ell_x) = \prod_{\sigma\in\mathscr{X}(L/K)} \sigma(x),$$

d'où le résultat. □

Nous continuons en démontrant un résultat fort utile pour les calculs, et qui généralise les résultats du lemme XXVII-4.4.

4.6. Proposition. *Soit L/K une extension de degré fini, et soit F/K une sous-extension. Alors, pour tout $x \in L$, on a*

$$\mathrm{Tr}_{L/K}(x) = \mathrm{Tr}_{F/K}\big(\mathrm{Tr}_{L/F}(x)\big) \quad \textit{et } N_{L/K}(x) = N_{F/K}\big(N_{L/F}(x)\big).$$

Démonstration. Gardons les notations de l'énoncé.

On commence par démontrer la proposition dans le cas particulier où $x \in F$. La proposition XXVII-4.3 montre que cela revient à établir les égalités

$$\mathrm{Tr}_{L/K}(x) = [L : F]\cdot \mathrm{Tr}_{F/K}(x) \quad \text{et } N_{L/K}(x) = N_{F/K}(x)^{[L:F]}.$$

Considérons pour cela $\mathbf{e} = (e_1, \ldots, e_n)$ une F-base de L et $\mathbf{f} = (f_1, \ldots, f_m)$ une K-base de F. Alors, la famille

$$\mathbf{g} = (e_i f_j)_{i,j} = (e_1 f_1, \ldots, e_1 f_m, \ldots, e_n f_1, \ldots, e_n f_m)$$

est une K-base de L. Si $\ell_{x,L}$ et $\ell_{x,F}$ désignent respectivement les endomorphismes de multiplication à gauche par x dans L et dans F, on a alors

$$\mathrm{Mat}(\ell_{x,L}\,;\mathbf{g}) = \begin{pmatrix} \mathrm{Mat}(\ell_{x,F}\,;\mathbf{f}) & & \\ & \ddots & \\ & & \mathrm{Mat}(\ell_{x,F}\,;\mathbf{f}) \end{pmatrix}.$$

Les égalités désirées en découle immédiatement.

Démontrons maintenant le résultat voulu dans un second cas particulier, à savoir lorsque $L = F(x)$.

Soit $n = [F(x) : F]$, et écrivons

$$\mu_{x,F} = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0,\ a_i \in F.$$

Soit $\mathbf{f} = (f_1, \ldots, f_m)$ une K-base de F, et soit $A_i = \mathrm{Mat}(\ell_{a_i,F}\,;\mathbf{f})$.

Puisque $L = F(x)$, d'après le lemme XXVII-4.4, on a

$$\mathrm{Tr}_{L/F}(x) = -a_{n-1} \quad \text{et} \quad N_{L/F}(x) = (-1)^n a_0,$$

si bien que

$$\mathrm{Tr}_{F/K}\big(\mathrm{Tr}_{L/F}(x)\big) \;=\; -\mathrm{Tr}_{F/K}(a_{n-1}) \;=\; -\mathrm{tr}(A_{n-1}),$$

ainsi que

$$N_{F/K}\big(N_{L/F}(x)\big) \;=\; (-1)^{mn} N_{F/K}(a_0) \;=\; (-1)^{mn}\det(A_0).$$

Mais, la matrice de $\ell_{x,L}$ dans la base $(f_1, \ldots, f_m, \ldots, f_1 x^{n-1}, \ldots, f_m x^{n-1})$ est la matrice par blocs

$$M = \begin{pmatrix} 0 & 0 & \cdots & 0 & -A_0 \\ \mathrm{I}_m & 0 & \cdots & 0 & -A_1 \\ 0 & \mathrm{I}_m & \cdots & 0 & -A_2 \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \mathrm{I}_m & -A_{n-1} \end{pmatrix}.$$

On a alors $\mathrm{Tr}_{L/K}(x) = \mathrm{tr}(M) = -\mathrm{tr}(A_{n-1}) = \mathrm{Tr}_{F/K}\big(\mathrm{Tr}_{L/F}(x)\big)$.

Par ailleurs, en décalant m fois les colonnes de M vers la droite, on obtient

$$N_{L/K}(x) \;=\; \det(M) \;=\; \big((-1)^{mn-1}\big)^m \begin{vmatrix} -A_0 & 0 & \cdots & 0 \\ -A_1 & \mathrm{I}_m & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ -A_{n-1} & 0 & \cdots & \mathrm{I}_m \end{vmatrix},$$

c'est-à-dire

$$N_{L/K}(x) \;=\; (-1)^{m(mn-1)} \det(-A_0) \;=\; (-1)^{mn} \det(A_0),$$

soit $N_{L/F}(x) = N_{F/K}\big(N_{L/F}(x)\big)$.

La proposition est donc démontrée dans le cas où $L = F(x)$.

Revenons au cas général. Soit $x \in L$. D'après le premier point, appliqué à la sous-extension $F(x)/K$, on a

$$\mathrm{Tr}_{L/K}(x) = [L : F(x)]\cdot \mathrm{Tr}_{F(x)/K}(x) \quad \text{et} \quad N_{L/F}(x) = N_{F(x)/K}(x)^{[L:F(x)]}.$$

D'après le second point, on a alors

$$\mathrm{Tr}_{L/K}(x) \;=\; [L : F(x)]\cdot \mathrm{Tr}_{F/K}\big(\mathrm{Tr}_{F(x)/F}(x)\big),$$

ainsi que

$$N_{L/K}(x) \;=\; N_{F/K}\big(N_{F(x)/F}(x)\big)^{[L:F(x)]},$$

soit encore

$$\mathrm{Tr}_{L/K}(x) \;=\; \mathrm{Tr}_{F/K}\big([L : F(x)]\cdot \mathrm{Tr}_{F(x)/F}(x)\big),$$

ainsi que

$$N_{L/K}(x) \;=\; N_{F/K}\big(N_{F(x)/F}^{[L:F(x)]}\big).$$

Une nouvelle application du premier point (ou du lemme XXVII-4.4) donne alors le résultat voulu. □

4.7. Remarque. Soit $x \in L$. En appliquant la proposition à $F = K(x)$, on retrouve les résultats du lemme XXVII-4.4.

Nous allons conclure ce chapitre par une caractérisation des extensions séparables en termes de trace. Cela nécessite de connaître la définition d'une forme bilinéaire non dégénérée, ainsi que sa caractérisation en termes de matrice représentative. Le lecteur se reportera au chapitre XXXII pour plus de détails.

Nous commençons par un lemme calculatoire qui sera utile pour la suite, ainsi que dans les exercices.

4.8. Lemme. *Soit L/K une extension de corps séparable de degré $n \geqslant 1$, et soient $\sigma_1 = \mathrm{Id}_L, \ldots, \sigma_n$ les n plongements de L/K.*

Pour tous $x_1, \dots, x_n \in L$, on pose

$$D_{L/K}(x_1, \dots, x_n) = \det\Big(\big(\mathrm{Tr}_{L/K}(x_i x_j)\big)_{i,j}\Big).$$

Alors, on a les propriétés suivantes :

(1) *pour tous $x_1, \dots, x_n \in L$, on a*

$$\big(\mathrm{Tr}_{L/K}(x_i x_j)\big)_{i,j} = M^t M,$$

avec $M = \big(\sigma_i(x_j)\big)_{i,j} \in \mathrm{M}_n(\overline{K})$. En particulier, on a

$$D_{L/K}(x_1, \dots, x_n) = \det(M)^2 ;$$

(2) *soient $x_1, y_1 \dots, x_n, y_n \in K$. On suppose qu'il existe $P \in \mathrm{M}_n(K)$ tel que* $\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = P \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$. *Alors, on a*

$$D_{L/K}(y_1, \dots, y_n) = \det(P)^2 D_{L/K}(x_1, \dots, x_n) ;$$

(3) *soit $\alpha \in L$ un élément primitif de L, et soit $f = \mu_{\alpha,K}$. Alors, on a*

$$D_{L/K}(1, \alpha, \dots, \alpha^{n-1}) = \mathrm{disc}(f) = (-1)^{\frac{n(n-1)}{2}} N_{L/K}\big(f'(\alpha)\big).$$

Démonstration. Gardons les notations du lemme. Le terme de $M^t M$ en position (i,j) est égal à

$$\sum_{k=1}^{n} \sigma_k(x_i)\sigma_k(x_j) = \sum_{k=1}^{n} \sigma_k(x_i x_j) = \mathrm{Tr}_{L/K}(x_i x_j),$$

d'où (1).

Montrons (2). La linéarité des plongements de L/K implique aisément l'égalité

$$\big(\sigma_i(y_j)\big)_{i,j} = \big(\sigma_i(x_j)\big)_{i,j} P^t.$$

On conclut alors en appliquant le point précédent.

Soit maintenant $\alpha \in L$ un élément primitif de L. On a donc $L = K(\alpha)$, et le polynôme $f = \mu_{\alpha,K} \in K[X]$ est un polynôme séparable de degré n.

Pour tout entier $i \in [\![1, n]\!]$, posons $\alpha_i = \sigma_i(\alpha)$. Ainsi, $\alpha = \alpha_1, \dots, \alpha_n \in \overline{K}$ sont les n racines distinctes de f. On a alors

$$M = \big(\sigma_i(\alpha^{j-1})\big)_{i,j} = \big(\sigma_i(\alpha)^{j-1}\big)_{i,j} = (\alpha_i^{j-1})_{i,j}.$$

La matrice M étant une matrice de Vandermonde, on obtient

$$\det(M) = \prod_{i<j} (\alpha_j - \alpha_i).$$

Le point (1) entraîne alors que l'on a l'égalité

$$D_{L/K}(1,\alpha,\dots,\alpha^{n-1}) = \prod_{i<j}(\alpha_j-\alpha_i)^2.$$

Cette dernière quantité est égale à $\mathrm{disc}(f)$, le polynôme f étant unitaire. Montrons l'égalité manquante. Puisque $f = \prod_{k=1}^{n}(X-\alpha_k)$, on obtient

$$f' = \sum_{j=1}^{n}\prod_{k\neq j}(X-\alpha_k).$$

Par conséquent, pour tout $j \in [\![1,n]\!]$, on a

$$\sigma_j\big(f'(\alpha)\big) = f'\big(\sigma_j(\alpha)\big) = f'(\alpha_j) = \prod_{k\neq j}(\alpha_j-\alpha_k).$$

D'après la proposition précédente, on a alors

$$N_{L/K}\big(f'(\alpha)\big) = \prod_{i=1}^{n}\sigma_i\big(f'(\alpha)\big) = \prod_{i\neq j}(\alpha_j-\alpha_i).$$

Remarquons maintenant que l'on a

$$\prod_{i<j}(\alpha_j-\alpha_i)^2 = (-1)^{\frac{n(n-1)}{2}}\prod_{i\neq j}(\alpha_j-\alpha_i).$$

Cela achève la démonstration. □

On peut maintenant montrer le résultat suivant.

4.9. Théorème. *Soit L/K une extension de degré fini. Les propriétés suivantes sont équivalentes :*

(1) *l'extension L/K est séparable ;*

(2) *la forme bilinéaire symétrique*

$$\begin{aligned} b_L : L\times L &\longrightarrow K \\ (x,y) &\longmapsto \mathrm{Tr}_{L/K}(xy) \end{aligned}$$

est non dégénérée ;

(3) *la forme linéaire $\mathrm{Tr}_{L/K}$ est non nulle.*

Démonstration

(1) $\Longrightarrow$ (2). On doit démontrer que b_L possède une matrice représentative inversible. Nous allons proposer deux démonstrations de ce fait.

Première démonstration. Soit $n = [L:K]$, et soient $\sigma_1,\dots,\sigma_n \in \mathscr{X}(L/K)$ les n plongements de L/K. Soit $\mathbf{e} = (e_1,\dots,e_n)$ une K-base de L arbitraire. Posons

$$M = \big(\sigma_i(e_j)\big)_{i,j} \in \mathrm{M}_n(\overline{K}).$$

Par le lemme précédent, on a $\mathrm{Mat}(b_L,\mathbf{e}) = M^tM$.

Soient $L_1, \dots, L_n$ les lignes de M, et soient $\lambda_1, \dots, \lambda_n \in \overline{K}$ tels que

$$\lambda_1 L_1 + \cdots + \lambda_n L_n = 0.$$

On a donc

$$\lambda_1 \sigma_1(e_j) + \cdots + \lambda_n \sigma_n(e_j) = 0, \quad \text{pour tout } j \in [\![1, n]\!].$$

Comme $\mathbf{e}$ est une K-base de L, on obtient

$$\lambda_1 \sigma_1 + \cdots + \lambda_n \sigma_n = 0.$$

Par le corollaire 1.5, on en déduit que $\lambda_i = 0$, pour tout $i \in [\![1, n]\!]$. Ainsi, les lignes de M sont indépendantes, et M est inversible. Mais alors, la matrice M^t est inversible, et $\mathrm{Mat}(b_L, \mathbf{e})$ l'est donc également.

Seconde démonstration. Puisque L/K est séparable, il existe un élément primitif $\alpha \in L$. Ainsi, $L = K(\alpha)$, et $\mathbf{e}_\alpha = (1, \alpha, \dots, \alpha^{n-1})$ est une K-base de L. D'après le lemme 4.8 (3), on a

$$\det\big(\mathrm{Mat}(b_L, \mathbf{e}_\alpha)\big) = \mathrm{disc}(\mu_{\alpha,K}) \in K^\times.$$

Ainsi, $\mathrm{Mat}(b_L, \mathbf{e})$ est inversible.

Par conséquent, b_L est non dégénérée.

(2) $\Longrightarrow$ (3). Si $\mathrm{Tr}_{L/K} = 0$, b_L est nulle, donc dégénérée, d'où le résultat par contraposition.

(3) $\Longrightarrow$ (1). Nous allons encore raisonner par contraposition.

Supposons que L/K ne soit pas séparable. Par le lemme 3.5 (2), K est de caractéristique $p > 0$. Soit M/K la sous-extension de L/K formée des éléments de L séparables sur K. Si L/M était séparable, L/K serait séparable par la proposition 2.5. Ainsi, il existe $y \in L$ non séparable sur M. On a donc $\mu_{y,M} = Q(X^p)$, où $Q \in M[X]$ est non constant, toujours d'après le lemme 3.5 (2). En particulier, on a $p \mid [M(y) : M]$, et par conséquent $p \mid [L : M]$.

Cela étant dit, soit $x \in L$. Rappelons que l'on a

$$\mathrm{Tr}_{L/K}(x) = [L : K(x)] \cdot \mathrm{Tr}_{K(x)/K}(x)$$

par le lemme 4.4.

Si x est séparable sur K, alors $x \in M$. On a donc $K(x) \subset M$, et en particulier, on a $[L : K(x)] = [L : M][M : K(x)]$. Ainsi, $p \mid [L : K(x)]$. Puisque K est de caractéristique p, on a alors $\mathrm{Tr}_{L/K}(x) = 0$.

Si x n'est pas séparable sur K, on a $\mu_{x,K} = P(X^p)$, où $P \in K[X]$ est de degré $d \geqslant 1$. Ainsi, $\mu_{x,K}$ est de degré dp, mais ne possède pas de terme en X^{dp-1}. Le lemme 4.4 montre alors que $\mathrm{Tr}_{K(x)/K}(x) = 0$, et l'on a encore $\mathrm{Tr}_{L/K}(x) = 0$. Autrement dit, $\mathrm{Tr}_{L/K}$ est identiquement nulle. Cela achève la démonstration. □

5. Exercices

Dans les exercices, on se fixe une clôture algébrique $\overline{K}$ de K. Toutes les extensions algébriques considérées seront des sous-extensions de $\overline{K}/K$. En particulier, on considérera tous les composita dans $\overline{K}/K$.

E1. Soient L/K et L'/K deux extensions algébriques linéairement disjointes. Montrer que l'application

$$\begin{aligned} \mathscr{X}(LL'/K) &\longrightarrow \mathscr{X}(L/K) \times \mathscr{X}(L'/K) \\ \sigma &\longmapsto (\sigma_{|L}, \sigma_{|L'}) \end{aligned}$$

est bijective.

E2. Soit K un corps. On dit que L/K est une clôture séparable de K si l'extension L/K est séparable, et si tout polynôme $P \in L[X]$ **séparable** possède au moins une racine dans L.

Montrer que tout corps possède une clôture séparable, unique à isomorphisme de K-algèbres près.

E3. Calculer $\mathscr{X}(L/K)$ et $\mathrm{Gal}(L/K)$ dans chacun des cas suivants.

(i) $K = \mathbb{Q}$, $L = \mathbb{Q}(\sqrt[4]{2})$;

(ii) $K = \mathbb{F}_2(T)$, $L = K(\alpha)$, où $\alpha \in \overline{K}$ est une racine de $X^4 + TX^2 + T$.

(iii) $K = \mathbb{Q}$, $L = \mathbb{Q}(\sqrt{2}, \sqrt{3}, \sqrt{5})$.

E4. Soit K un corps de caractéristique différente de 2. Soient $a, b \in K^\times$ tels que a, b, $ab \notin K^{\times 2}$, et soit $L = K(\sqrt{a}, \sqrt{b})$.

Montrer que $[L : K] = 4$, que L/K est séparable, et calculer $\mathscr{X}(L/K)$.

E5. Soit L/K une extension de degré fini. Montrer que si $\mathrm{car}(K) \nmid [L : K]$, alors L/K est séparable.

E6. Soit L/K une extension algébrique. Le but de cet exercice est de démontrer que L/K admet un nombre fini de sous-extensions si, et seulement si, L/K est monogène.

a. On suppose que $L = K(\alpha)$.

(i) Soit F/K une sous-extension de L/K. Montrer qu'il existe un facteur unitaire P de $\mu_{\alpha,K}$ dans $L[X]$ tel que F/K est engendrée par les coefficients de P.

(ii) En déduire que L/K n'a qu'un nombre fini de sous-extensions.

(iii) Le résultat est-il encore vrai si L/K n'est pas algébrique ?

b. On suppose que L/K n'a qu'un nombre fini de sous-extensions.

(i) Montrer que $[L : K]$ est fini.

(ii) Montrer que si K est fini, L/K est monogène.

(iii) On suppose que K est infini. Montrer que pour tous $\alpha, \beta \in L$, il existe $\lambda \in K$ tel que $K(\alpha, \beta) = K(\alpha + \lambda\beta)$.

(iv) En déduire que L/K est monogène.

On se propose maintenant d'exhiber une extension de degré fini possédant un nombre infini de sous-extensions.

Soit p un nombre premier, soit $K = \mathbb{F}_p(X^p, Y^p)$ et soit $L = \mathbb{F}_p(X, Y)$. Pour tout $k \geqslant 1$, on pose

$$L_k = K(X + X^{p^k} Y).$$

c. Montrer que $[L : K] = p^2$, et que $[L_k : K] = p$ pour tout $k \geqslant 1$.

d. Montrer que les sous-extensions L_k/K sont deux à deux distinctes, et donc que L/K est une extension de degré fini possédant une infinité de sous-extensions.

E7. Soit K un corps de caractéristique $p > 0$.

a. Montrer que tout polynôme irréductible unitaire $P \in K[X]$ s'écrit de manière unique sous la forme

$$P = Q(X^{p^r}), \quad Q \in K[X],\ r \geqslant 0,$$

où Q est irréductible unitaire et séparable sur K.

b. Montrer qu'un polynôme irréductible unitaire $P \in K[X]$ a une unique racine dans $\overline{K}$ si, et seulement si $P = X^{p^r} - a$, avec $r \geqslant 0$ et $a \in K$.

On dit que $\alpha \in \overline{K}$ est *purement inséparable* sur K s'il existe $m \geqslant 0$ tel que $\alpha^{p^m} \in K$. On dit qu'une extension L/K est purement inséparable si L/K est algébrique et tous ses éléments sont purement inséparables sur K.

c. Soit $\alpha \in \overline{K}$ un élément purement inséparable. Montrer que

$$\mu_{\alpha,K} = X^{p^r} - \alpha^{p^r},$$

où $r \geqslant 0$ est le plus petit entier tel que $\alpha^{p^r} \in K$.

d. Soit L/K une extension de degré fini. Montrer l'équivalence des propriétés suivantes :

(1) tout morphisme d'anneaux $\iota : K \longrightarrow \overline{K}$ possède une unique extension à L ;

(2) $|\mathscr{X}(L/K)| = 1$;

(3) l'extension L/K est purement inséparable.

e. Montrer qu'une extension purement inséparable de degré fini est de degré une puissance de p.

f. Soit L/K une extension. Montrer que l'ensemble M des éléments purement inséparables sur K est une sous-extension de L. Montrer que si $L = \overline{K}$, M est un corps parfait.

g. Soit L/K une extension de corps. Montrer qu'il existe une unique sous-extension M/K telle que M/K soit séparable et L/M soit purement inséparable.

h. Soit $K = \mathbb{F}_2(T_1, T_2)$, et soit $P = X^4 + T_1X^2 + T_2 \in K[X]$.

(i) Montrer que P est irréductible.

Soit $L = K(\alpha)$, où $\alpha \in \overline{K}$ est une racine de P.

(ii) Montrer par l'absurde qu'il n'existe pas d'extension M/K purement inséparable telle que L/M soit séparable.

E8. Soit K un corps fini, et soit L/K une extension de degré fini. Montrer que la trace et la norme sont surjectives.

E9. Dans ce qui suit, K est un corps quelconque, $n \geqslant 2$, et $a \in K^\times$. Le but de cet exercice est d'étudier l'irréductibilité du polynôme $X^n - a$ dans $K[X]$.

a. Montrer que les degrés des facteurs irréductibles de $X^n - a$ sont divisibles par l'ordre de $\overline{a} \in K^\times/K^{\times n}$. En déduire que si $X^n - a$ admet un facteur irréductible premier à n, alors $a \in K^{\times n}$.

Indication. On pourra poser $X^n - a = PQ$, avec $P, Q \in K[X]$ non constants, et montrer que $(-1)^n \deg(P) P(0)^n = a^{\deg(P)}$.

On suppose maintenant que $n = p^r$, où p est nombre premier et $r \geqslant 1$.

b. Montrer que si $a \in K^{\times p}$, alors $X^{p^r} - a$ n'est pas irréductible. On distinguera les cas $\mathrm{car}(K) = p$ et $\mathrm{car}(K) \neq p$.

c. On suppose que $p \nmid \mathrm{car}(K)$ et que $a \in K^\times \setminus K^{\times p}$.

(i) Montrer que si $\beta \in \overline{K}$ est une racine de $X^p - a$, on a une factorisation

$$X^{p^r} - a = \prod_{\sigma \in \mathscr{X}(K(\beta)/K)} \left(X^{p^{r-1}} - \sigma(\beta)\right).$$

(ii) Montrer que $N_{K(\beta)/K}(\beta) = (-1)^{p+1} a$.

(iii) On distinguera alors les cas $p = 2$ et p impair et l'on raisonnera par récurrence sur r pour étudier l'irréductibilité de $X^{p^r} - a$. Pour le cas $p = 2$, on remarquera que

$$X^4 + 4b^4 = (X^2 + 2bX + 2b^2)(X^2 - 2bX + 2b^2).$$

On revient au cas général.

d. Montrer que $X^n - a$ est irréductible dans $K[X]$ si, et seulement si, les deux conditions suivantes sont vérifiées :

(1) pour tout p premier divisant n, on a $a \in K^\times \setminus K^{\times p}$;

(2) si $4 \mid n$, on a $a \in K^\times \setminus -4K^{\times 2}$.

Indication. On pourra raisonner par récurrence sur le nombre de diviseurs premiers de n.

E10. Soit $L/\mathbb{Q}$ une extension de degré n, et soit $\mathcal{O}_L$ l'anneau des éléments de L entiers sur $\mathbb{Z}$. Le but de cet exercice est de montrer que $\mathcal{O}_L$ est un $\mathbb{Z}$-module libre de rang n et un anneau noethérien.

a. Montrer que pour tout $x \in \mathcal{O}_L$, on a

$$\mathrm{Tr}_{L/\mathbb{Q}}(x), \quad N_{L/\mathbb{Q}}(x) \in \mathbb{Z}.$$

b. Montrer que si $x \in L$, il existe $\lambda \in \mathbb{Z}$ tel que $\lambda x \in \mathcal{O}_L$.

c. En déduire l'existence d'une $\mathbb{Q}$-base $(e_1, \ldots, e_n)$ de L formée d'éléments de $\mathcal{O}_L$. Soit $(e_1^*, \ldots, e_n^*)$ la base duale par rapport à b. Montrer que l'on a

$$\mathcal{O}_L \subset \mathbb{Z}e_1^* \oplus \ldots \oplus \mathbb{Z}e_n^*.$$

d. En déduire que $\mathcal{O}_L$ est un $\mathbb{Z}$-module libre de rang n, ainsi qu'un anneau noethérien.

E11. Soit $L/\mathbb{Q}$ une extension de degré n. Si $x_1, \ldots, x_n \in L$, on rappelle que

$$D_{L/\mathbb{Q}}(x_1, \ldots, x_n) = \det\left(\mathrm{Tr}_{L/\mathbb{Q}}(x_i x_j)\right)_{i,j}.$$

a. Montrer que $D_{L/\mathbb{Q}}(x_1, \ldots, x_n) \in \mathbb{Q}^\times$ pour toute $\mathbb{Q}$-base $(x_1, \ldots, x_n)$ de L.

b. Soit $(\omega_1, \ldots, \omega_n)$ une $\mathbb{Z}$-base[8] de $\mathcal{O}_L$.

En utilisant le lemme 4.8, montrer que $D_{L/\mathbb{Q}}(\omega_1, \ldots, \omega_n)$ est un entier non nul, et qu'il ne dépend pas du choix de la base $(\omega_1, \ldots, \omega_n)$. Cet entier est noté d_L, et est appelé le *discriminant* de $L/\mathbb{Q}$.

c. Soient $x_1, \ldots, x_n \in \mathcal{O}_L$. Montrer qu'il existe $m \geqslant 0$ tel que

$$D_{L/\mathbb{Q}}(x_1, \ldots, x_n) = m^2 d_L,$$

et que l'on a $D_{L/\mathbb{Q}}(x_1, \ldots, x_n) = d_L$ si, et seulement si, $(x_1, \ldots, x_n)$ est une $\mathbb{Z}$-base de $\mathcal{O}_L$.

d. Utiliser l'exercice XXVI-14 pour calculer d_L lorsque $L = \mathbb{Q}(\sqrt{d})$.

8. Une telle base existe d'après l'exercice 10.

E12. Soit $L/\mathbb{Q}$ une extension de degré fini. En séparant les permutations en deux classes à gauche modulo $\mathfrak{A}_n$ dans l'expression du déterminant, démontrer que l'on a $d_L \equiv 0$ ou 1 [4].

E13. Soit $L/\mathbb{Q}$ une extension de degré fini, et soit $\alpha \in \mathcal{O}_L$ tel que $L = \mathbb{Q}(\alpha)$. Soit $f = \mu_{\alpha,\mathbb{Q}}$.

a. En utilisant le théorème de la base adaptée, montrer que $[\mathcal{O}_L : \mathbb{Z}[\alpha]]$ est fini et que
$$\operatorname{disc}(f) = [\mathcal{O}_L : \mathbb{Z}[\alpha]]^2 d_L.$$

b. En déduire que $\mathcal{O}_L = \mathbb{Z}[\alpha]$ dans chacun des cas suivants :
(i) $\operatorname{disc}(f)$ est sans facteur carré ;
(ii) $\operatorname{disc}(f) = 4m$, où m est sans facteur carré et $m \not\equiv 1 \mod 4$.

E14. Soit $L = \mathbb{Q}(\alpha)$, où $\alpha \in \mathbb{C}$ est une racine du polynôme f de $\mathbb{Z}[X]$ donné par $f(X) = X^3 - X^2 - 2X - 8$.

a. Montrer que f est irréductible sur $\mathbb{Q}$.

b. Soit $\beta = \dfrac{\alpha^2 + \alpha}{2}$. Montrer que $\beta^3 - 3\beta^2 - 10\beta - 8 = 0$, et en déduire que β est entier sur $\mathbb{Z}$.

c. Vérifier que $D_{L/\mathbb{Q}}(1, \alpha, \beta) = -503$.

d. En déduire que la famille $(1, \alpha, \beta)$ est une $\mathbb{Z}$-base de $\mathcal{O}_L$.

e. Montrer que $D_{L/\mathbb{Q}}(1, \theta, \theta^2)$ est pair pour tout $\theta \in \mathcal{O}_L$. En déduire que l'on a $\mathcal{O}_L \neq \mathbb{Z}[\theta]$ pour tout $\theta \in \mathcal{O}_L$ [9].

Indication. Écrire $\theta = a + b\alpha + c\beta$, $a, b, c \in \mathbb{Z}$, et calculer $D_{L/\mathbb{Q}}(1, \theta, \theta^2)$ modulo 2 en fonction a, b et c.

E15. Soit $L/\mathbb{Q}$ une extension de degré n, et soit $\alpha \in \mathcal{O}_L$ tel que $L = \mathbb{Q}(\alpha)$. Soit $f = \mu_{\alpha,\mathbb{Q}}$. On suppose que f est un polynôme d'Eisenstein en un nombre premier p. Autrement dit,
$$f = X^n + pm_{n-1}X^{n-1} + \cdots + pm_1X + pm_0,$$
avec $m_i \in \mathbb{Z}$ et $p \nmid m_0$. On veut montrer que $p \nmid [\mathcal{O}_L : \mathbb{Z}[\alpha]]$.

On suppose que $p \mid [\mathcal{O}_L : \mathbb{Z}[\alpha]]$.

a. Justifier l'existence de $\beta \in \mathcal{O}_L$ tel que $\beta \notin \mathbb{Z}[\alpha]$ et $p\beta \in \mathbb{Z}[\alpha]$.

b. En déduire qu'il existe $\gamma \in \mathcal{O}_L \setminus \mathbb{Z}[\alpha]$ tel que
$$p\gamma = \alpha^r(a_0 + a_1\alpha + \cdots + a_{n-r-1}\alpha^{n-r-1}),$$
pour un certain entier $r \in [\![0, n-1]\!]$, et des entiers $a_i \in \mathbb{Z}$, avec $p \nmid a_0$.

9. Merci à Olivier Bordellès de m'avoir proposé cet exercice.

c. En remarquant que $\alpha^n \in p\mathcal{O}_L$, en déduire que $a_0\alpha^{n-1} \in p\mathcal{O}_L$.

d. En calculant la norme, aboutir à une contradiction.

e. On suppose que pour tout p premier tel que $p^2 \mid \mathrm{disc}(f)$, le polynôme f est un polynôme d'Eisenstein en p. Montrer que $\mathcal{O}_L = \mathbb{Z}[\alpha]$.

f. Soit $L = \mathbb{Q}(\sqrt[3]{2})$. Déterminer $\mathcal{O}_L$.

g. Soit p un nombre premier, et soit $L = \mathbb{Q}(\zeta_p)$. Calculer $\mathrm{disc}(\Phi_p)$. En considérant $\zeta_p - 1$, montrer que $\mathcal{O}_L = \mathbb{Z}[\zeta_p]$.

Chapitre XXVIII

Extensions galoisiennes

Nous nous intéressons plus spécifiquement dans ce chapitre au groupe des automorphismes d'une extension de degré fini. Dans la suite, nous identifierons $\mathrm{Gal}(L/K)$ avec l'image de l'application

$$\mathrm{Gal}(L/K) \longrightarrow \mathscr{X}(L/K),$$

obtenue en composant à droite par l'inclusion $L \subset \overline{K}$. Comme nous l'avons déjà remarqué, on a alors une inclusion

$$\mathrm{Gal}(L/K) \subset \mathscr{X}(L/K),$$

mais l'inclusion est stricte en général.

Nous allons en premier lieu nous intéresser aux conditions sous lesquelles cette inclusion est une égalité.

1. Extensions normales

Entrons immédiatement dans le vif du sujet.

1.1. Lemme. *Soit L/K une extension de degré fini. Les propriétés suivantes sont équivalentes :*

(1) *tout plongement de L/K est un automorphisme de L/K ;*

(2) *tout polynôme irréductible unitaire de $K[X]$ ayant au moins une racine dans L est scindé dans L ;*

(3) *l'extension L/K est le corps des racines d'un polynôme de $K[X]$.*

Démonstration

(1) $\Longrightarrow$ (2). Soit $P \in K[X]$ un polynôme irréductible unitaire possédant une racine $\alpha \in L$. Soient $\alpha_1 = \alpha, \alpha_2, \ldots, \alpha_n \in \overline{K}$ les racines distinctes de P dans $\overline{K}$. On a donc $\mu_{\alpha,K} = P$.

Pour tout $i \in [\![1, n]\!]$, soit $\tau_i : K(\alpha) \longrightarrow \overline{K}$ l'unique plongement de $K(\alpha)/K$ tel que $\tau_i(\alpha) = \alpha_i$. Soit alors $\sigma_i : L \longrightarrow \overline{K}$ une extension de τ_i à L. Par hypothèse, σ_i est un automorphisme de L/K. En particulier, $\sigma_i(L) = L$. Mais alors, on a

$$\sigma_i(\alpha) = \tau_i(\alpha) = \alpha_i \in L,$$

et P est scindé dans L.

(2) $\Longrightarrow$ (3). Soient $\alpha_1, \ldots, \alpha_m \in L$ tels que $L = K(\alpha_1, \ldots, \alpha_m)$. Par hypothèse, $\mu_{\alpha_i,K}$ est scindé dans L pour tout $i \in [\![1, m]\!]$. En particulier, L contient toutes les racines du polynôme

$$P = \prod_{i=1}^{m} \mu_{\alpha_i,K}.$$

Comme l'ensemble des racines P contient $\alpha_1, \ldots, \alpha_m$, on en déduit que l'extension L/K est engendrée par les racines de P. Autrement dit, L/K est le corps des racines de P.

(3) $\Longrightarrow$ (1). Supposons que l'extension L/K soit le corps des racines d'un polynôme $P \in K[X]$, que l'on peut supposer unitaire. Il existe donc des éléments $\alpha_1, \ldots, \alpha_m \in L$ tels que

$$P = (X - \alpha_1) \cdots (X - \alpha_m) \quad \text{et} \quad L = K(\alpha_1, \ldots, \alpha_m).$$

Soit $\sigma \in \mathscr{X}(L/K)$. Puisque σ est K-linéaire, on a

$$P\big(\sigma(\alpha_i)\big) = \sigma\big(P(\alpha_i)\big) = 0,$$

et $\sigma(\alpha_i)$ est donc une racine de P. Ainsi, on a

$$\sigma(\alpha_i) \in \{\alpha_1, \ldots, \alpha_m\} \subset L.$$

Or, puisque $L = K[\alpha_1, \ldots, \alpha_m]$, on en déduit que $\sigma(L) \subset L$. Ainsi, σ est donc un morphisme de K-algèbres injectif de L dans lui-même. Comme L/K est de degré fini, on en déduit que σ est bijectif. Par conséquent, σ est un automorphisme de L/K.

Cela achève la démonstration. □

1.2. Définition. Une extension L/K de degré fini est dite *normale* si elle vérifie une des conditions équivalentes du lemme précédent.

1.3. Exemples

(1) Soit $K = \mathbb{Q}$ et soit $L = \mathbb{Q}(\sqrt[3]{2})$. Alors, l'extension L/K n'est pas normale. En effet,

$$P = X^3 - 2 = (X - \sqrt[3]{2})(X - j\sqrt[3]{2})(X - j^2\sqrt[3]{2})$$

est irréductible dans $\mathbb{Q}[X]$ et a une racine dans L, à savoir $\sqrt[3]{2}$, mais n'a pas toutes ses racines dans L, puisque $L \subset \mathbb{R}$ et P a des racines complexes non réelles.

(2) Soit $K = \mathbb{Q}$ et soit $L = \mathbb{Q}(\sqrt[3]{2}, j)$. Alors, L/K est normale. En effet, on a les égalités

$$L = \mathbb{Q}(\sqrt[3]{2}, j, j^2) = \mathbb{Q}(\sqrt[3]{2}, j\sqrt[3]{2}, j^2\sqrt[3]{2}).$$

En particulier, L/K est le corps des racines de $X^3 - 2 \in K[X]$.

(3) Soit $K = \mathbb{F}_2(T^2)$ et $L = \mathbb{F}_2(T)$. Alors, L/K est le corps des racines du polynôme $X^2 - T^2 \in K[X]$. En effet, on a

$$X^2 - T^2 = (X - T)^2.$$

Ainsi, L/K est normale.

(4) Soit $K = \mathbb{Q}$ et soit $L = \mathbb{Q}(\zeta_n)$, avec $n \geqslant 1$. Les racines complexes du polynôme $X^n - 1$ étant les ζ_n^k, avec $k \in [\![1, n]\!]$, on en déduit que L/K est le corps des racines de $X^n - 1$. Par conséquent, l'extension L/K est normale.

(5) Puisque $\mathbb{F}_q$ est le corps des racines de $X^q - X$ sur $\mathbb{F}_p$, l'extension $\mathbb{F}_q/\mathbb{F}_p$ est normale.

Attention ! Une extension normale n'est pas nécessairement le corps des racines d'un polynôme **irréductible** (cf. exercice XXVI-3).

1.4. Lemme. *Soient L/K et L'/K deux extensions normales. Alors, LL'/K est normale.*

Démonstration. Si L/K et L'/K sont respectivement les corps de racines de P et Q respectivement, on vérifie facilement que LL'/K est le corps des racines de PQ, d'où le résultat. □

1.5. Lemme. *Soit L/K une extension, et soient E/K et E'/K deux sous-extensions. Si E/K est normale, alors EE'/E' est normale.*

En particulier, si L/K est normale, et si M/K est une sous-extension, alors L/M est normale.

Démonstration. Si E/K est le corps des racines de $P \in K[X]$, alors EE'/E' est le corps des racines de P, vu comme polynôme de $E'[X]$. La seconde partie s'obtient en appliquant ce qui précède à $E = L$ et $E' = M$. □

1.6. Remarque. Les exemples 1.3 (1) et (2) montrent qu'une sous-extension d'une extension normale n'est pas nécessairement normale.

1.7. Lemme. *L'intersection d'une famille non vide d'extensions normales de K est une extension normale de K.*

Démonstration. Soit $(L_i/K)_{i\in I}$ une famille non vide d'extensions normales, et soit

$$L = \bigcap_{i\in I} L_i.$$

Alors, L/K est de degré fini. Soit $P \in K[X]$ un polynôme irréductible unitaire ayant une racine dans L. Comme $L \subset L_i$ pour tout $i \in I$, l'hypothèse sur L_i/K montre que toutes les racines de P dans $\overline{K}$ sont en fait dans L_i, et cela pour tout $i \in I$. Autrement dit, les racines de P dans $\overline{K}$ sont dans L, d'où le résultat. □

Pour finir ce paragraphe, remarquons qu'une extension L/K de degré fini est contenue dans une extension normale. En effet, si $L = K(\alpha_1, \ldots, \alpha_m)$, alors L est contenue dans le corps des racines de $\mu_{\alpha_1,K} \cdots \mu_{\alpha_m,K}$. Cela justifie la définition suivante.

1.8. Définition. Soit L/K une extension de degré fini. La *clôture normale* de L/K est l'intersection de tous les extensions normales contenant L. Autrement dit, c'est la plus petite sous-extension normale contenant L/K.

2. Extensions galoisiennes

Soit L/K une extension de degré fini. Rappelons que l'on a l'inclusion

$$\mathrm{Gal}(L/K) \subset \mathscr{X}(L/K),$$

et donc

$$|\mathrm{Gal}(L/K)| \leqslant |\mathscr{X}(L/K)| \leqslant [L : K],$$

par le lemme XXVII-1.6. Nous allons maintenant déterminer sous quelles conditions on a $|\mathrm{Gal}(L/K)| = [L : K]$.

2.1. Lemme. *Soit L/K une extension de corps de degré fini. Alors, on a l'égalité $|\mathrm{Gal}(L/K)| = [L : K]$ si, et seulement si, L/K est normale et séparable.*

Démonstration. Si $|\mathrm{Gal}(L : K)| = [L : K]$, on a donc $|\mathscr{X}(L/K)| = [L : K]$, ce qui prouve que L/K est séparable d'après le lemme XXVII-1.6, ainsi que l'égalité $\mathrm{Gal}(L/K) = \mathscr{X}(L/K)$, ce qui montre que L/K est normale par définition.

Inversement, supposons que L/K soit normale et séparable. La normalité de L/K implique alors que $|\mathrm{Gal}(L/K)| = |\mathscr{X}(L/K)|$, et la séparabilité montre alors que $|\mathscr{X}(L/K)| = [L : K]$, d'où l'égalité requise. Cela achève la démonstration. □

Cela motive la définition suivante.

2.2. Définition. Une extension L/K de degré fini est dite *galoisienne* si elle est normale et séparable. Dans ce cas, on a

$$\mathrm{Gal}(L/K) = [L : K].$$

C'est même équivalent par le lemme précédent.

2.3. Exemples

(1) Soit $n \geqslant 1$. Alors, l'extension $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ est galoisienne. En effet, elle est normale par l'exemple 1.3 (4) et séparable, car $\mathbb{Q}$ est de caractéristique nulle (cf. théorème XXVII-3.7).

On peut aussi remarquer que, d'après le lemme XXV-4.11, on a

$$\mathrm{Gal}\big(\mathbb{Q}(\zeta_n)/\mathbb{Q}\big) \simeq (\mathbb{Z}/n\mathbb{Z})^\times.$$

En particulier, on a

$$|\mathrm{Gal}\big(\mathbb{Q}(\zeta_n)/\mathbb{Q}\big)| = \varphi(n) = [\mathbb{Q}(\zeta_n) : \mathbb{Q}],$$

ce qui redémontre que cette extension est galoisienne.

(2) Soit q une puissance d'un nombre premier p et soit $d \geqslant 1$. Alors, l'extension $\mathbb{F}_{q^d}/\mathbb{F}_q$ est galoisienne.

En effet, l'extension $\mathbb{F}_{q^d}/\mathbb{F}_p$ est normale par l'exemple 1.3 (5). On en déduit que $\mathbb{F}_{q^d}/\mathbb{F}_q$ est normale par le lemme 1.5. De plus, cette extension est séparable par le théorème XXVII-3.7.

On peut retrouver ce fait en remarquant que $\mathrm{Gal}(\mathbb{F}_{q^d}/\mathbb{F}_q)$ est cyclique d'ordre d par la proposition XXVI-2.11. En particulier, on a

$$|\mathrm{Gal}(\mathbb{F}_{q^d}/\mathbb{F}_q)| = d = [\mathbb{F}_{q^d} : \mathbb{F}_q].$$

(3) Si $K = \mathbb{F}_2(T^2)$ et $L = \mathbb{F}_2(T)$, l'extension est normale, non séparable d'après les exemples 1.3 (3) et 1.9 (3).

(4) Soit $K = \mathbb{F}_2(T)$, et soit $L = K(\sqrt[3]{T}, \sqrt{T})$. Alors, L/K n'est ni normale, ni séparable (cf. exercice 2).

(5) L'extension $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$ est séparable, car $\mathbb{Q}$ est de caractéristique nulle, mais n'est pas normale d'après l'exemple 1.3 (1).

(6) L'extension $\mathbb{Q}(\sqrt[3]{2}, j)/\mathbb{Q}$ est galoisienne. En effet, elle est normale par l'exemple 1.3 (2), et séparable, car $\mathbb{Q}$ est de caractéristique nulle. Calculons son groupe de Galois. Il revient au même de calculer les plongements de cette extension.

Puisque $\mu_{\sqrt[3]{2},\mathbb{Q}} = X^3 - 2$, les plongements de $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$ sont les morphismes

$$\tau_k : \sqrt[3]{2} \rightsquigarrow j^k \sqrt[3]{2}, \quad k \in [\![0, 2]\!].$$

On remarque que $j, j^2 \notin \mathbb{Q}(\sqrt[3]{2})$, puisque $\mathbb{Q}(\sqrt[3]{2}) \subset \mathbb{R}$. On a donc

$$\mu_{j,\mathbb{Q}(\sqrt[3]{2})} = X^2 + X + 1,$$

puisque que ce polynôme annule j et n'a pas de racines dans $\mathbb{Q}(\sqrt[3]{2})$. Ainsi, on a $\tau_k(\mu_{j,\mathbb{Q}(\sqrt[3]{2})}) = \mu_{j,\mathbb{Q}(\sqrt[3]{2})}$ pour tout $k \in [\![0, 2]\!]$.

On en déduit alors que les éléments de $\mathrm{Gal}\big(\mathbb{Q}(\sqrt[3]{2}, j)/\mathbb{Q}\big)$ sont les morphismes

$$\sigma_{k,m} : \left| \begin{array}{ccc} \sqrt[3]{2} & \rightsquigarrow & j^k \sqrt[3]{2} \\ j & \rightsquigarrow & j^m \end{array} \right. , \quad k \in [\![0, 2]\!], \ m \in [\![1, 2]\!].$$

Soient $\sigma = \sigma_{1,1}$ et $\tau = \sigma_{0,2}$. On vérifie aisément que σ et τ ne commutent pas. Ainsi, $\mathrm{Gal}\big(\mathbb{Q}(\sqrt[3]{2}, j)/\mathbb{Q}\big)$ est un groupe d'ordre 6 non abélien, d'où un isomorphisme

$$\mathrm{Gal}\big(\mathbb{Q}(\sqrt[3]{2}, j)/\mathbb{Q}\big) \simeq D_3.$$

De simples calculs montrent d'ailleurs que σ est d'ordre 3, que τ est d'ordre 2, et que $\tau\sigma\tau^{-1} = \sigma^{-1}$.

2.4. Lemme. *Toute extension séparable de degré fini est contenue dans une extension galoisienne.*

Démonstration. Soit F/K une extension séparable de degré fini. D'après le théorème de l'élément primitif, il existe $\alpha \in L$ tel que $F = K(\alpha)$. Soit L/K le corps des racines de $\mu_{\alpha,K}$. Par hypothèse, α est séparable sur K et le polynôme $\mu_{\alpha,K} \in K[X]$ est donc séparable. En particulier, toutes les racines de $\mu_{\alpha,K}$ sont aussi séparables sur K. Comme L/K est engendrée par ces racines, L/K est séparable sur K. De plus, L/K est normale par définition. Ainsi, L/K est une extension galoisienne. De plus, $\alpha \in L$, donc $K(\alpha) \subset L$, d'où le résultat. □

Passons aux propriétés élémentaires des extensions galoisiennes.

2.5. Proposition. *Soit L/K une extension, et soient E/K et E'/K deux sous-extensions. Si E/K est galoisienne, alors EE'/E' est galoisienne.*

De plus, pour tout $\sigma \in \mathrm{Gal}(EE'/E')$, *on a* $\sigma_{|E} \in \mathrm{Gal}(E/K)$, *et l'application*

$$\begin{array}{ccc} \mathrm{Gal}(EE'/E') & \longrightarrow & \mathrm{Gal}(E/K) \\ \sigma & \longmapsto & \sigma_{|E} \end{array}$$

est un morphisme de groupes injectif, identifiant ainsi $\mathrm{Gal}(EE'/E')$ *à un sous-groupe de* $\mathrm{Gal}(E/K)$.

En particulier, si L/K *est galoisienne, et si* M/K *est une sous-extension, alors* L/M *est galoisienne et* $\mathrm{Gal}(L/M)$ *est un sous-groupe de* $\mathrm{Gal}(L/K)$.

Démonstration. D'après le corollaire XXVII-2.3, EE'/E' est séparable, et d'après le lemme 1.5, EE'/E' est normale. Ainsi, EE'/E' est galoisienne. Soit maintenant $\sigma \in \mathrm{Gal}(EE'/E')$. Comme σ est un morphisme de E'-algèbres, c'est aussi un morphisme de K-algèbres. Ainsi, $\sigma_{|E}$ est un plongement de E/K, donc un automorphisme de E/K par hypothèse. Il est alors clair que l'application

$$\begin{array}{ccc} \mathrm{Gal}(EE'/E') & \longrightarrow & \mathrm{Gal}(E/K) \\ \sigma & \longmapsto & \sigma_{|E} \end{array}$$

est un morphisme de groupes. Supposons maintenant que $\sigma \in \mathrm{Gal}(EE'/E')$ vérifie $\sigma_{|E} = \mathrm{Id}_E$. Le lemme XXV-3.20 et la E'-linéarité de σ impliquent alors que $\sigma = \mathrm{Id}_{EE'}$, d'où l'injectivité annoncée.

Le dernier point s'obtient en appliquant ce qui précède à $E = L$ et $E' = M$, auquel cas le morphisme précédent est en fait une inclusion (on peut aussi raisonner directement). Cela achève la démonstration. □

2.6. Lemme. *L'intersection d'une famille non vide d'extensions galoisiennes est galoisienne.*

Démonstration. Soit $(L_i/K)_{i\in I}$ une famille non vide d'extensions galoisiennes, et soit

$$L = \bigcap_{i\in I} L_i.$$

Alors, L/K est normale par le lemme 1.7. Par ailleurs, comme L/K est une sous-extension d'une extension séparable (par exemple, c'est une sous-extension de L_i/K, où $i \in I$ est fixé arbitrairement), L/K est aussi séparable par la remarque 1.10, d'où le résultat. □

Le lemme 2.4 et ce dernier lemme nous permet de donner la définition suivante.

2.7. Définition. Soit F/K une extension séparable de degré fini. L'intersection de toutes les extensions galoisiennes de degré fini contenant F est appelée la *clôture galoisienne* de F/K.

2.8. Exemple. Soit F/K une extension séparable de degré fini. Si $\alpha \in F$ est un élément primitif, la démonstration du lemme 2.4 montre que le corps des racines L/K de $\mu_{\alpha,K}$ est une extension galoisienne contenant F. C'est en fait la clôture galoisienne de F/K.

En effet, soit L'/K une extension galoisienne contenant F. Alors, le corps L' contient α, et $\mu_{\alpha,K}$ a donc une racine dans L'. Comme L'/K est normale, L' contient toutes les racines de $\mu_{\alpha,K}$, et contient par conséquent la sous-extension engendrée par ces racines, c'est-à-dire L.

2.9. Lemme. *Soient L_1/K et L_2/K deux extensions galoisiennes. Alors, l'extension L_1L_2/K est galoisienne. De plus, l'application*

$$\begin{array}{ccc} \mathrm{Gal}(L_1L_2/K) & \longrightarrow & \mathrm{Gal}(L_1/K) \times \mathrm{Gal}(L_2/K) \\ \sigma & \longmapsto & (\sigma_{|L_1}, \sigma_{|L_2}) \end{array}$$

est un morphisme de groupes injectif, et identifie donc $\mathrm{Gal}(L_1L_2/K)$ à un sous-groupe de $\mathrm{Gal}(L_1/K) \times \mathrm{Gal}(L_2/K)$.

Démonstration. Le premier point découle du corollaire XXVII-2.4 ainsi que du lemme 1.4.

Soit $\sigma \in \mathrm{Gal}(L_1L_2/K)$. Pour $i \in \{1,2\}$, $\sigma_{|L_i}$ est un plongement de L_i/K, donc un élément de $\mathrm{Gal}(L_i/K)$ puisque L_i/K est galoisienne. On vérifie alors aisément que l'application de l'énoncé est un morphisme de groupes. Supposons maintenant que $\sigma \in \mathrm{Gal}(L_1L_2/K)$ vérifie

$$\sigma_{|L_i} = \mathrm{Id}_{L_i}, \quad \text{pour tout } i \in \{1,2\},$$

le lemme XXV-3.20 montre alors que $\sigma = \mathrm{Id}_{L_1L_2}$, d'où le résultat. □

Profitons-en également pour énoncer un résultat annoncé sans démonstration dans le paragraphe XXV-5, et qui explique de manière plus conceptuelle le point (1) de la proposition XXV-4.12.

2.10. Lemme. *Soit L/K une extension galoisienne. Alors, $L^{\mathrm{Gal}(L/K)} = K$.*

Démonstration. Puisque L/K est galoisienne, on a $\mathscr{X}(L/K) = \mathrm{Gal}(L/K)$. Il suffit alors d'appliquer le corollaire XXVII-2.12 pour conclure. □

On continue ce paragraphe par une proposition qui facilite grandement les calculs.

2.11. Proposition. *Soient E/K et E'/K deux sous-extensions d'une extension L/K. Si E/K est galoisienne, alors E/K et E'/K sont linéairement disjointes si, et seulement si, $E \cap E' = K$.*

En particulier, si E/K est galoisienne, alors $E/E \cap E'$ et $E'/E \cap E'$ sont linéairement disjointes.

Démonstration. On sait déjà que la condition $E \cap E' = K$ est nécessaire. Supposons maintenant que $E \cap E' = K$. Supposons que E/K et E'/K ne soient pas linéairement disjointes. Soit $(x'_1, \ldots, x'_n)$ une famille K-libre de E' qui est E-liée. On suppose $n \geqslant 1$ minimal. Remarquons que $n \geqslant 2$. Quitte à renuméroter, on peut supposer qu'il existe $x_1, \ldots, x_{n-1} \in E$ tels que

$$x'_n = \sum_{i=1}^{n-1} x_i x'_i.$$

Pour tout $\sigma \in \mathrm{Gal}(EE'/E')$, on a donc $x'_n = \sum_{i=1}^{n-1} \sigma(x_i) x'_i$, et par conséquent

$$\sum_{i=1}^{n-1} (\sigma(x_i) - x_i) x'_i = 0, \qquad \text{pour tout } \sigma \in \mathrm{Gal}(EE'/E').$$

Par la proposition 2.5, on a $\sigma(x_i) \in E$, et par suite $\sigma(x_i) - x_i \in E$ pour tout $i \in [\![1, n-1]\!]$. Par minimalité de n, on en déduit

$$\sigma(x_i) = x_i, \qquad \text{pour tout } i \in [\![1, n-1]\!].$$

Puisque l'extension EE'/E' est galoisienne, le lemme 2.10 implique alors que $x_i \in E'$ pour tout $i \in [\![1, n-1]\!]$. Comme on a aussi $x_i \in E$, on obtient que $x_i \in K$ pour tout $i \in [\![1, n-1]\!]$. Mais alors, on en déduit que la famille $(x'_1, \ldots, x'_n)$ est K-liée, d'où une contradiction. Ainsi, E/K et E'/K sont bien linéairement disjointes.

Le dernière point provient du fait que $E/E \cap E'$ est aussi galoisienne d'après la proposition 2.5, et du point précédent. Cela achève la démonstration. □

Signalons maintenant deux applications de cette proposition. La première précise les résultats de la proposition 2.5.

2.12. Théorème. *Soit L/K une extension, et soient E/K et E'/K deux sous-extensions. Si E/K est galoisienne, alors l'extension EE'/E' est galoisienne, et pour tout $\sigma \in \mathrm{Gal}(EE'/E')$, on a $\sigma_{|E} \in \mathrm{Gal}(E/E \cap E')$, et l'application*

$$\begin{array}{rccc} \rho : & \mathrm{Gal}(EE'/E') & \longrightarrow & \mathrm{Gal}(E/E \cap E') \\ & \sigma & \longmapsto & \sigma_{|E} \end{array}$$

est un isomorphisme de groupes.

En particulier, on a $\mathrm{Gal}(EE'/E') \simeq \mathrm{Gal}(E/K)$ si, et seulement si, E/K et E'/K sont linéairement disjointes.

Démonstration. Puisque E/K est galoisienne, $E/E \cap E'$ est aussi galoisienne par la dernière partie de la proposition 2.5. En appliquant la première partie de cette proposition aux extensions $E/E \cap E'$ et $E'/E \cap E'$,

on obtient que pour tout $\sigma \in \mathrm{Gal}(EE'/E')$, on a $\sigma_{|E} \in \mathrm{Gal}(E/E \cap E')$, et que l'application

$$\begin{array}{rccc} \rho : & \mathrm{Gal}(EE'/E') & \longrightarrow & \mathrm{Gal}(E/E \cap E') \\ & \sigma & \longmapsto & \sigma_{|E} \end{array}$$

est un morphisme de groupes injectif. Comme $E/E \cap E'$ et $E'/E \cap E'$ sont linéairement disjointes d'après la proposition précédente, on a

$$[EE' : E'] = [E : E \cap E'].$$

Mais alors, $|\mathrm{Gal}(EE'/E')| = |\mathrm{Gal}(E/E \cap E')|$, et ρ est donc bijective.

Montrons la dernière partie. Si E/K et E'/K sont linéairement disjointes, alors $E \cap E' = K$, et le premier point fournit un isomorphisme de groupes

$$\mathrm{Gal}(EE'/E') \simeq \mathrm{Gal}(E/K).$$

Supposons enfin que les groupes $\mathrm{Gal}(EE'/E')$ et $\mathrm{Gal}(E/K)$ soient isomorphes. En particulier, ils ont même ordre, et donc $[EE' : E'] = [E : K]$. Mais alors, E/K et E'/K sont linéairement disjointes, et cela d'après la proposition XXV-3.25. Cela achève la démonstration. □

2.13. Théorème. *Soient L_1/K et L_2/K deux extensions galoisiennes. Alors, les extensions $L_1/L_1 \cap L_2, L_2/L_1 \cap L_2$ et $L_1L_2/L_1 \cap L_2$ sont galoisiennes, et l'application*

$$\begin{array}{rccc} \theta : & \mathrm{Gal}(L_1L_2/L_1 \cap L_2) & \longrightarrow & \mathrm{Gal}(L_1/L_1 \cap L_2) \times \mathrm{Gal}(L_2/L_1 \cap L_2) \\ & \sigma & \longmapsto & (\sigma_{|L_1}, \sigma_{|L_2}) \end{array}$$

est un isomorphisme de groupes.

En particulier, on a $\mathrm{Gal}(L_1L_2/K) \simeq \mathrm{Gal}(L_1/K) \times \mathrm{Gal}(L_2/K)$ si, et seulement si, L_1/K et L_2/K sont linéairement disjointes.

Démonstration. Soit $i \in \{1, 2\}$. Puisque L_i/K est galoisienne, $L_i/L_1 \cap L_2$ est galoisienne par la proposition 2.5. Le lemme 2.9 montre alors que l'extension $L_1L_2/L_1 \cap L_2$ est galoisienne.

Pour tout $\sigma \in \mathrm{Gal}(L_1L_2/L_1 \cap L_2)$, $\sigma_{|L_i}$ est un plongement de $L_i/L_1 \cap L_2$, donc un élément de $\mathrm{Gal}(L_i/L_1 \cap L_2)$ puisque $L_i/L_1 \cap L_2$ est galoisienne. On vérifie aisément que θ est un morphisme de groupes. De plus, les deux extensions $L_1/L_1 \cap L_2$ et $L_2/L_1 \cap L_2$ sont linéairement disjointes par la proposition 2.11. On a donc

$$[L_1L_2 : L_1 \cap L_2] = [L_1 : L_1 \cap L_2][L_2 : L_1 \cap L_2]$$

par le corollaire XXV-3.26. Comme les extensions $L_1/L_1 \cap L_2$, $L_2/L_1 \cap L_2$ et $L_1L_2/L_1 \cap L_2$ sont galoisiennes, cela revient à dire que les deux groupes finis $\mathrm{Gal}(L_1L_2/L_1 \cap L_2)$ et $\mathrm{Gal}(L_1/L_1 \cap L_2) \times \mathrm{Gal}(L_2/L_1 \cap L_2)$ ont même ordre.

Il suffit donc de montrer que θ est injectif pour en conclure que c'est un isomorphisme. Or, si $\sigma \in \mathrm{Gal}(L_1L_2/L_1 \cap L_2)$ vérifie

$$\sigma_{|L_i} = \mathrm{Id}_{L_i}, \quad \text{pour tout } i \in \{1,2\},$$

le lemme XXV-3.20 montre alors que $\sigma = \mathrm{Id}_{L_1L_2}$, d'où le premier point.

Notons que L_1L_2/K est galoisienne par le lemme 2.9. Si les extensions L_1/K et L_2/K sont linéairement disjointes, on a $L_1 \cap L_2 = K$, et le point précédent donne un isomorphisme de groupes

$$\mathrm{Gal}(L_1L_2/K) \simeq \mathrm{Gal}(L_1/K) \times \mathrm{Gal}(L_2/K).$$

Inversement, si $\mathrm{Gal}(L_1L_2/K)$ et $\mathrm{Gal}(L_1/K) \times \mathrm{Gal}(L_2/K)$ sont isomorphes, ils ont même ordre, et par conséquent

$$[L_1L_2 : K] = [L_1 : K] \times [L_2 : K],$$

puisque les extensions $L_1/K, L_2/K$ et L_1L_2/K sont galoisiennes. Par le corollaire XXV-3.26, L_1/K et L_2/K sont linéairement disjointes. Cela achève la démonstration. □

3. Le groupe de Galois d'un polynôme

Dans la présente section, nous allons nous intéresser au groupe de Galois du corps des racines d'un polynôme. Introduisons tout d'abord quelques notations.

Notation. Si $P \in K[X]$ est un polynôme, on note $\mathrm{Dec}_K(P)$ le corps des racines de P dans $\overline{K}$. On note $\mathrm{Gal}_K(P)$ le groupe de Galois de $\mathrm{Dec}_K(P)$.

3.1. Définition. Soit $P \in K[X]$ un polynôme. Le *groupe de Galois* de P est le groupe de Galois de son corps des racines. On le note $\mathrm{Gal}_K(P)$.

3.2. Remarque

Si $P \in K[X]$ est un polynôme séparable, $\mathrm{Dec}_K(P)/K$ est une extension galoisienne. En effet, c'est une extension normale par définition. De plus, si $\alpha \in L$ est une racine de P, on a $\mu_{\alpha,K} \mid P$. Comme P n'a que des racines simples dans L, il en est de même de $\mu_{\alpha,K}$. Or, L/K est engendré par les racines de P, qui sont des éléments séparables, et L/K est donc séparable par le corollaire XXVII-2.2.

On peut maintenant donner une autre caractérisation des extensions galoisiennes.

3.3. Lemme. *Une extension L/K de degré fini est galoisienne si, et seulement si, L/K est le corps de décomposition d'un polynôme $P \in K[X]$ irréductible séparable.*

Dans ce cas, on peut choisir P de degré $[L : K]$.

Démonstration. Si $L = \mathrm{Dec}_K(P)$, avec $P \in K[X]$ irréductible séparable, alors L/K est galoisienne de degré fini d'après la remarque précédente. Inversement, soit L/K une extension galoisienne de degré fini. Comme L/K est séparable, le théorème de l'élément primitif assure l'existence de $\alpha \in L$ tel que $L = K(\alpha)$. Posons $P = \mu_{\alpha,K}$. Puisque, L/K est séparable, α est séparable, et ainsi P est un polynôme irréductible séparable. De plus, on a

$$[L : K] = [K(\alpha) : K] = \deg(P).$$

Montrons que $L = \mathrm{Dec}_K(P)$. Par construction, le polynôme P possède une racine dans L, à savoir α. Puisque L/K est normale, P possède toutes ses racines dans L. Ainsi, L contient toutes les racines de P, et par conséquent

$$\mathrm{Dec}_K(P) \subset L.$$

Par ailleurs, on a

$$L = K(\alpha) \subset \mathrm{Dec}_K(P),$$

puisque α est une racine de P. Finalement, on obtient $L = \mathrm{Dec}_K(P)$. Cela achève la démonstration. □

3.4. Remarque. Si $P \in K[X]$ est un polynôme irréductible séparable de degré $n \geqslant 1$, il se peut très bien que l'on ait $[\mathrm{Dec}_K(P) : K] > n$.

Par exemple, si $P = X^3 - 2 \in \mathbb{Q}[X]$, l'exemple 1.3 (2) montre que l'on a

$$\mathrm{Dec}_{\mathbb{Q}}(P) = \mathbb{Q}(\sqrt[3]{2}, j),$$

qui est de degré 6 sur $\mathbb{Q}$ d'après l'exemple 2.3 (6).

Cela démontre également qu'une extension galoisienne de degré n peut être égale au corps de décomposition d'un polynôme irréductible séparable de degré strictement inférieur à n.

Nous allons maintenant étudier de plus près le groupe de Galois d'un polynôme séparable.

3.5. Lemme. *Soit $P \in K[X]$ un polynôme de degré $n \geqslant 1$. Alors :*

(1) *le groupe $\mathrm{Gal}_K(P)$ stabilise l'ensemble des racines dans $\overline{K}$ de tout facteur irréductible de P ;*

(2) *si X_P est l'ensemble des racines de P dans $\overline{K}$, l'application*

$$\begin{aligned} \mathrm{Gal}_K(P) \times X_P &\longrightarrow X_P \\ (\sigma, \alpha) &\longmapsto \sigma(\alpha) \end{aligned}$$

définit une action fidèle de $\mathrm{Gal}_K(P)$ sur X_P.
Autrement dit, l'application

$$\begin{aligned} \mathrm{Gal}_K(P) &\longrightarrow \mathfrak{S}(X_P) \\ \sigma &\longmapsto \sigma_{|X_P} \end{aligned}$$

est un morphisme de groupes injectif. En particulier, si P est séparable, on a $[\mathrm{Dec}_K(P) : K] \mid n!$;

(3) *si P est irréductible, $\mathrm{Gal}_K(P)$ agit transitivement sur l'ensemble des racines de P. Si de plus P est séparable, la réciproque est vraie.*

Démonstration

(1) Soit $Q \in K[X]$ un facteur irréductible unitaire de P. Alors, pour toute racine $\alpha \in \mathrm{Dec}_K(P)$ de Q, et tout $\sigma \in \mathrm{Gal}_K(P)$, on a

$$Q\big(\sigma(\alpha)\big) = \sigma\big(Q(\alpha)\big) = 0.$$

Ainsi, $\sigma(\alpha)$ est encore une racine de Q.

(2) Le point précédent montre en particulier que tout élément de $\mathrm{Gal}_K(P)$ stabilise X_P (fait déjà établi et utilisé maintes fois). Le fait que l'application

$$\begin{aligned} \mathrm{Gal}_K(P) \times X_P &\longrightarrow X_P \\ (\sigma, \alpha) &\longmapsto \sigma(\alpha) \end{aligned}$$

définisse une action de $\mathrm{Gal}_K(P)$ sur X_P est clair. Montrons que cette action est fidèle. Si $\sigma \in \mathrm{Gal}_K(P)$ fixe chaque racine $\alpha_1, \ldots, \alpha_n \in \overline{K}$ de P, alors σ fixe chaque élément de $\mathrm{Dec}_K(P)$, puisque

$$\mathrm{Dec}_K(P) = K(\alpha_1, \ldots, \alpha_n) = K[\alpha_1, \ldots, \alpha_n].$$

Ainsi, σ est le morphisme identité, et l'action est bien fidèle. Comme le morphisme correspondant à cette action n'est rien d'autre que le morphisme

$$\begin{aligned} \mathrm{Gal}_K(P) &\longrightarrow \mathfrak{S}(X_P) \\ \sigma &\longmapsto \sigma_{|X_P}, \end{aligned}$$

celui-ci est injectif. En particulier, le groupe $\mathrm{Gal}_K(P)$ s'identifie à un sous-groupe de $\mathfrak{S}(X_P)$. Si P est séparable, alors $|X_P| = n$, et de plus $\mathrm{Dec}_K(P)/K$ est galoisienne. On a alors

$$[\mathrm{Dec}_K(P) : K] = |\mathrm{Gal}_K(P)| \mid |\mathfrak{S}(X_P)|,$$

ce qui est exactement le résultat souhaité.

(3) Supposons que P soit irréductible. On peut toujours supposer P unitaire. Soient α, $\beta \in \mathrm{Dec}_K(P)$ deux racines de P. Puisque P est irréductible unitaire, on a $\mu_{\alpha,K} = P$. D'après le théorème XXVI-2.2, il existe un plongement $\sigma_\beta : K(\alpha) \longrightarrow \overline{K}$ tel que $\sigma_\beta(\alpha) = \beta$. Ce plongement se prolonge alors en un plongement $\sigma : \mathrm{Dec}_K(P) \longrightarrow \overline{K}$. Comme $\mathrm{Dec}_K(P)/K$ est une extension normale, $\sigma \in \mathrm{Gal}_K(P)$. Par construction, $\sigma(\alpha) = \beta$, et l'action de $\mathrm{Gal}_K(P)$ est bien transitive.

Supposons maintenant que P soit séparable. Si P n'est pas irréductible, il possède au moins deux facteurs irréductibles unitaires distincts Q_1 et Q_2. Pour $i = 1, 2$, soit X_i l'ensemble des racines de Q_i dans $\mathrm{Dec}_K(P)$. Par hypothèse sur P, ces deux sous-ensembles sont disjoints. Par (1), tout élément de $\mathrm{Gal}_K(P)$ stabilise X_i. On en déduit que l'orbite de toute racine de Q_i est contenue dans X_i. En particulier, les orbites d'une racine de Q_1 et d'une racine de Q_2 sont distinctes, et l'action de $\mathrm{Gal}_K(P)$ n'est pas transitive.

Cela achève la démonstration. □

3.6. Remarque. Le point (2) montre en particulier que, pour tout polynôme séparable $P \in K[X]$ de degré n, on a $[\mathrm{Dec}_K(P) : K] \leqslant n!$. La remarque 3.4 entraîne que cette borne ne peut être améliorée en général.

Le lemme précédent montre que l'on a un morphisme injectif

$$\begin{array}{ccc} \mathrm{Gal}_K(P) & \longrightarrow & \mathfrak{S}(X_P) \\ \sigma & \longmapsto & \sigma_{|X_P}. \end{array}$$

Nous allons expliciter cette image lorsque K est un corps fini.

3.7. Exemples. Soit $K = \mathbb{F}_q$, et soit $P \in K[X]$ un polynôme unitaire séparable. Alors, $\mathrm{Dec}_K(P)$ est un corps fini. Posons

$$\begin{array}{ccc} \sigma_q : \mathrm{Dec}_K(P) & \longrightarrow & \mathrm{Dec}_K(P) \\ x & \longmapsto & x^q. \end{array}$$

(1) Supposons que P soit irréductible de degré $d \geqslant 1$.
Alors, $\mathrm{Gal}_K(P)$ est cyclique d'ordre d, engendré par σ_q, d'après la proposition XXVI-2.11. D'après la remarque XXVII-2.13 par exemple, si $\alpha \in \mathrm{Dec}_K(P)$ est une racine de P fixée, les racines de $\mu_{\alpha,K} = P$ sont

$$\alpha,\ \sigma_q(\alpha),\ \ldots,\ \sigma_q^{d-1}(\alpha).$$

En particulier, σ_q induit un d-cycle sur l'ensemble X des racines de P. Ainsi, $\mathrm{Gal}_K(P)$ s'identifie dans $\mathfrak{S}(X)$ au sous-groupe cyclique engendré par un d-cycle.

(2) Supposons maintenant que $P = P_1 \cdots P_r$, où $P_1, \ldots, P_r$ sont irréductibles unitaires et deux à deux distincts. Alors, le groupe $\mathrm{Gal}_K(P)$ est toujours engendré par σ_q. De plus, d'après le lemme 3.5 et le point précédent, la restriction de σ_q à l'ensemble X_i des racines de P_i est un d_i-cycle, où $d_i = \deg(P_i)$. Comme P est séparable, $X_1, \ldots, X_r$ forment une partition de X, et l'on en déduit que tous ces cycles sont à supports deux à deux disjoints. Par conséquent, σ_q induit une permutation de X de type $(d_1, \ldots, d_r)$.

On continue maintenant par un résultat qui détermine sous quelles conditions l'image de ce morphisme est incluse dans $\mathfrak{A}(X_P)$.

3.8. Théorème. *Soit $P \in K[X]$ un polynôme séparable de degré $n \geqslant 1$. Alors, l'extension $K\big(\sqrt{\mathrm{disc}(P)}\big)/K$ est une sous-extension de $\mathrm{Dec}_K(P)/K$. De plus, si X_P est l'ensemble des racines de P dans $\overline{K}$, l'image du morphisme injectif*

$$\begin{array}{ccc} \mathrm{Gal}_K(P) & \longrightarrow & \mathfrak{S}(X_P) \\ \sigma & \longmapsto & \sigma_{|X_P} \end{array}$$

est incluse dans $\mathfrak{A}(X_P)$ si, et seulement si, $\mathrm{disc}(P) \in K^{\times 2}$.

Démonstration. Soit $a_n \in K^\times$ le coefficient dominant de P, et soient

$$\alpha_1, \ldots, \alpha_n \in \mathrm{Dec}_K(P)$$

les racines de P. Par définition, l'élément

$$\gamma = a_n^{n-1} \prod_{i<j} (\alpha_j - \alpha_i) \in \mathrm{Dec}_K(P)$$

vérifie $\gamma^2 = \mathrm{disc}(P)$. On a donc bien $K\big(\sqrt{\mathrm{disc}(P)}\big) \subset \mathrm{Dec}_K(P)$.

Pour tout $\sigma \in \mathrm{Gal}_K(P)$, posons $M_\sigma = \big(\sigma(\alpha_j)^{i-1}\big)_{i,j}$. Cette matrice étant une matrice de Vandermonde, on a

$$\sigma(\gamma) = a_n^{n-1} \prod_{i<j} \big(\sigma(\alpha_j) - \sigma(\alpha_i)\big) = a_n^{n-1} \det(M_\sigma)$$

pour tout $\sigma \in \mathrm{Gal}_K(P)$. Remarquons que pour tout $\sigma \in \mathrm{Gal}_K(P)$, la permutation $\sigma_{|X_P}$ induit une permutation des colonnes de M_{Id}. La matrice obtenue n'étant rien d'autre que M_σ, on en déduit l'égalité

$$\det(M_\sigma) = \varepsilon(\sigma_{|X_P}) \det(M_{\mathrm{Id}}), \quad \text{pour tout } \sigma \in \mathrm{Gal}_K(P).$$

Autrement dit, on a

$$\sigma(\gamma) = \varepsilon(\sigma_{|X_P})\gamma, \quad \text{pour tout } \sigma \in \mathrm{Gal}_K(P).$$

Comme $\gamma \neq 0$ (puisque P est séparable), l'image du morphisme injectif

$$\begin{aligned} \mathrm{Gal}_K(P) &\longrightarrow \mathfrak{S}(X_P) \\ \sigma &\longmapsto \sigma_{|X_P} \end{aligned}$$

est donc incluse dans $\mathfrak{A}(X_P)$ si, et seulement si, γ est fixe par tout élément de $\mathrm{Gal}_K(P)$. D'après le lemme 2.10, cela est équivalent à $\gamma \in K^\times$ Cela revient à dire que $\mathrm{disc}(P) \in K^{\times 2}$. Cela achève la démonstration. □

3.9. Remarque. Soit $P \in K[X]$ un polynôme **irréductible** séparable de degré n. Après le choix d'une numérotation des racines de P, les résultats précédents montrent que le groupe $\mathrm{Gal}_K(P)$ s'identifie à un sous-groupe transitif de $\mathfrak{S}_n$ (i.e. un sous-groupe agissant transitivement sur $[\![1, n]\!]$), voire même à un sous-groupe transitif de $\mathfrak{A}_n$ si $\mathrm{disc}(P) \in K^{\times 2}$.

Ce simple fait limite donc les possibilités pour $\mathrm{Gal}_K(P)$ (même si le nombre de sous-groupes transitifs de $\mathfrak{S}_n$ peut être très grand). Par exemple, si P est un polynôme irréductible séparable de degré 5, il n'y a que cinq possibilités à isomorphisme près, à savoir $\mathfrak{S}_5$, $\mathfrak{A}_5$, $\mathbb{Z}/5\mathbb{Z}$, D_5 ou $\mathrm{GLA}(\mathbb{F}_5)$, d'après l'exercice XI-13.

Nous allons maintenant expliquer comment la réduction modulo p peut permettre d'obtenir des renseignements sur le groupe de Galois d'un polynôme unitaire $P \in \mathbb{Z}[X]$.

3.10. Théorème. *Soit $P \in \mathbb{Z}[X]$ un polynôme unitaire, soit p un nombre premier et soit $\overline{P} \in \mathbb{F}_p[X]$ sa réduction modulo p. On suppose que $\overline{P}$ est séparable. Alors, il existe un morphisme de groupes injectif*

$$\mathrm{Gal}_{\mathbb{F}_p}(\overline{P}) \hookrightarrow \mathrm{Gal}_{\mathbb{Q}}(P).$$

De plus, si $d_1, \ldots, d_r$ sont les degrés des facteurs irréductibles unitaires de $\overline{P}$, alors $\mathrm{Gal}_{\mathbb{Q}}(P)$ contient une permutation de type $(d_1, \ldots, d_r)$.

Démonstration. Écrivons $P = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0 \in \mathbb{Z}[X]$. D'après la proposition XXVI-1.11, il existe donc $h \in \mathbb{Z}[X_0, \ldots, X_{n-1}]$ tel que

$$\mathrm{disc}(P) = h(a_0, \ldots, a_{n-1}).$$

On a alors

$$\overline{\mathrm{disc}(P)} = \overline{h(a_0, \ldots, a_{n-1})} = h(\overline{a}_0, \ldots, \overline{a}_{n-1}) = \mathrm{disc}(\overline{P}).$$

Comme $\overline{P}$ est séparable, on a $\mathrm{disc}(\overline{P}) \neq \overline{0}$ par cette même proposition. En particulier, $\mathrm{disc}(P) \neq 0$ et P est séparable.

Soient $\alpha_1, \ldots, \alpha_n \in \mathrm{Dec}_K(P)$ les racines de P, et soit $A = \mathbb{Z}[\alpha_1, \ldots, \alpha_n]$. D'après le lemme de Krull, il existe un idéal maximal $\mathfrak{m}$ de A contenant l'idéal (p).

Posons $K = A/\mathfrak{m}$. C'est un corps de caractéristique de p, par définition même. Si $x \in A$, on notera $\overline{x}$ sa classe dans K.

Notons que la restriction de la projection canonique $\pi : A \longrightarrow K$ à $\mathbb{Z}$ induit un morphisme injectif $\mathbb{F}_p \longrightarrow K$. Dans l'anneau $K[X]$, on a alors les égalités

$$\overline{P} = \overline{(X-\alpha_1)\cdots(X-\alpha_n)} = (X-\overline{\alpha}_1)\cdots(X-\overline{\alpha}_n).$$

Cela démontre que $\overline{\alpha}_1, \ldots, \overline{\alpha}_n \in K$ sont les n racines distinctes de $\overline{P}$. Or, on a

$$K = \mathbb{F}_p[\overline{\alpha}_1, \ldots, \overline{\alpha}_n] = \mathbb{F}_p(\overline{\alpha}_1, \ldots, \overline{\alpha}_n),$$

la dernière égalité ayant lieu, car $\overline{\alpha}_1, \ldots, \overline{\alpha}_n$ sont algébriques sur $\mathbb{F}_p$. Il s'ensuit que $K = \mathrm{Dec}_{\mathbb{F}_p}(\overline{P})$. De plus, la projection canonique $\pi : A \longrightarrow K$ envoie bijectivement l'ensemble des racines de P dans $\mathrm{Dec}_{\mathbb{Q}}(P)$ sur l'ensemble des racines de $\overline{P}$ dans K.

Par définition, $\alpha_1, \ldots, \alpha_n$ sont entiers sur $\mathbb{Z}$. D'après le lemme XXVI-4.5, A est un groupe abélien de type fini, et il est sans torsion, car c'est un anneau intègre. C'est donc un groupe abélien libre, d'après le théorème XII-4.9 par exemple. Soit $(e_1, \ldots, e_m)$ une $\mathbb{Z}$-base de A.

Nous allons maintenant montrer que $m = [\mathrm{Dec}_{\mathbb{Q}}(P) : \mathbb{Q}]$. Tout d'abord,

$$\mathrm{Dec}_{\mathbb{Q}}(P) = \mathbb{Q}(\alpha_1, \ldots, \alpha_n) = \mathbb{Q}[\alpha_1, \ldots, \alpha_n].$$

On en déduit facilement que tout élément de $\mathrm{Dec}_{\mathbb{Q}}(P)$ s'écrit sous la forme

$$\frac{x}{k}, \quad \text{avec } x \in A \quad \text{et} \quad k \in \mathbb{Z} \setminus \{0\}.$$

En particulier, $(e_1, \ldots, e_m)$ est une famille $\mathbb{Q}$-génératrice de $\mathrm{Dec}_{\mathbb{Q}}(P)$. Or, la famille $(e_1, \ldots, e_m)$ étant $\mathbb{Z}$-libre, elle est aussi $\mathbb{Q}$-libre. Par conséquent, c'est une $\mathbb{Q}$-base de $\mathrm{Dec}_{\mathbb{Q}}(P)$, d'où l'égalité désirée.

Démontrons à présent qu'il y a au plus m morphismes d'anneaux $A \longrightarrow K$. Soient $\varphi_1, \ldots, \varphi_r : A \longrightarrow K$ des morphisme d'anneaux deux à deux distincts. Comme ces morphismes sont aussi des morphismes de monoïdes

$$(A, \cdot) \longrightarrow (K, \cdot),$$

le lemme XXVII-1.4 montre que la famille $(\varphi_1, \ldots, \varphi_r)$ est K-libre. Remarquons que, pour tous $\lambda_1, \ldots, \lambda_r \in K$, on a

$$\sum_{j=1}^{m} \lambda_j \varphi_j = 0 \iff \sum_{j=1}^{m} \lambda_j \varphi_j(e_i) = 0, \quad \text{pour tout } i \in [\![1, m]\!]$$

$$\iff \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_r \end{pmatrix} \in \mathrm{Ker}(M),$$

où $M = \big(\varphi_j(e_i)\big)_{i,j} \in \mathrm{M}_{m \times r}(K)$.

La famille $(\varphi_1, \ldots, \varphi_r)$ étant K-libre, on obtient $\mathrm{Ker}(M) = \{0\}$. En particulier, $r \leqslant m$.

La prochaine étape est de montrer qu'il y a en fait exactement m morphismes d'anneaux $A \longrightarrow K$. Remarquons pour cela que, puisque P est séparable, $\mathrm{Dec}_{\mathbb{Q}}(P)/\mathbb{Q}$ est galoisienne. Ainsi, $\mathrm{Gal}_{\mathbb{Q}}(P)$ possède m éléments. De plus, un élément $\sigma \in \mathrm{Gal}_{\mathbb{Q}}(P)$ envoie une racine de P sur une racine de P. En particulier, $\sigma(A) \subset A$, et l'on obtient un morphisme d'anneaux

$$\begin{aligned} \varphi_\sigma : A &\longrightarrow K \\ x &\longmapsto \overline{\sigma(x)}. \end{aligned}$$

Montrons que les m morphismes obtenus ainsi sont deux à deux distincts. Soient σ_1, $\sigma_2 \in \mathrm{Gal}_{\mathbb{Q}}(P)$ tels que $\varphi_{\sigma_1} = \varphi_{\sigma_2}$.

En particulier, on a

$$\overline{\sigma_1(\alpha_i)} = \overline{\sigma_2(\alpha_i)}, \quad \text{pour tout } i \in [\![1, n]\!].$$

Comme $\sigma_1(\alpha_i)$ et $\sigma_2(\alpha_i)$ sont des racines de P dans $\mathrm{Dec}_{\mathbb{Q}}(P)$, un point précédent montre que l'on a

$$\sigma_1(\alpha_i) = \sigma_2(\alpha_i), \quad \text{pour tout } i \in [\![1, n]\!].$$

On en déduit alors que $\sigma_1 = \sigma_2$, par le raisonnement habituel. Par conséquent, il y a exactement m morphismes d'anneaux $A \longrightarrow K$, à savoir les morphismes φ_σ lorsque σ parcourt $\mathrm{Gal}_{\mathbb{Q}}(P)$. Autrement dit, l'application

$$\begin{aligned} \mathrm{Gal}_{\mathbb{Q}}(P) &\longrightarrow \mathrm{Hom}(A, K) \\ \sigma &\longmapsto \varphi_\sigma \end{aligned}$$

est bijective.

Si maintenant $\tau \in \mathrm{Gal}_{\mathbb{F}_p}(\overline{P}) = \mathrm{Gal}(K/\mathbb{F}_p)$, $\tau \circ \pi$ est un morphisme d'anneaux $A \longrightarrow K$. Il existe donc un unique $\rho(\tau) \in \mathrm{Gal}_{\mathbb{Q}}(P)$ tel que

$$\tau \circ \pi = \varphi_{\rho(\tau)} = \pi \circ \rho(\tau)_{|A}.$$

Montrons que $\rho : \mathrm{Gal}_{\mathbb{F}_p}(\overline{P}) \longrightarrow \mathrm{Gal}_{\mathbb{Q}}(P)$ est un morphisme de groupes injectif. Pour tous τ, $\tau' \in \mathrm{Gal}_{\mathbb{F}_p}(\overline{P})$, on a

$$\begin{aligned} (\tau\tau') \circ \pi &= \tau \circ (\tau' \circ \pi) \\ &= \tau \circ \left(\pi \circ \rho(\tau')_{|A}\right) \\ &= (\tau \circ \pi) \circ \rho(\tau')_{|A} \\ &= \left(\pi \circ \rho(\tau)_{|A}\right) \circ \rho(\tau')_{|A} \\ &= \pi \circ \left(\rho(\tau) \circ \rho(\tau')\right)_{|A}. \end{aligned}$$

Par définition de $\rho(\tau\tau')$, on obtient l'égalité

$$\rho(\tau\tau') = \rho(\tau) \circ \rho(\tau').$$

Par conséquent, ρ est bien un morphisme de groupes. Soit $\tau \in \mathrm{Gal}_{\mathbb{F}_p}(\overline{P})$ tel

que $\rho(\tau) = \mathrm{Id}$. On a donc $\tau \circ \pi = \pi$. Comme π est surjective, d'image K, on obtient $\tau = \mathrm{Id}_K$, et ρ est donc injectif.

Il reste donc à démontrer le dernier point. Soit $\tau \in \mathrm{Gal}_{\mathbb{F}_p}(\overline{P})$. Notons alors s_τ, $t_\tau \in \mathfrak{S}_n$ les permutations de $\mathfrak{S}_n$ telles que

$$\tau(\overline{\alpha}_i) = \overline{\alpha}_{s_\tau(i)} \quad \text{et} \quad \rho(\tau)(\alpha_i) = \alpha_{t_\tau(i)}, \quad \text{pour tout } i \in [\![1, n]\!].$$

Par définition de ρ, on a $\overline{\alpha}_{s_\tau(i)} = \overline{\alpha}_{t_\tau(i)}$, et donc $\alpha_{s_\tau(i)} = \alpha_{t_\tau(i)}$ pour tout entier $i \in [\![1, n]\!]$. On a donc $s_\tau = t_\tau$. Il suffit donc de vérifier que le groupe $\mathrm{Gal}_{\mathbb{F}_p}(\overline{P})$ contient une permutation de type $(d_1, \ldots, d_r)$ pour conclure. C'est le cas d'après l'exemple 3.7. Cela achève la démonstration.□

3.11. Exemple. Soit $P = X^5 - X - 1 \in \mathbb{Q}[X]$.

D'après le petit théorème de Fermat, P n'a pas de racines dans $\mathbb{F}_5$, et est donc irréductible[1] modulo 5 par l'exercice XXVI-7. Ainsi, $\mathrm{Gal}_{\mathbb{Q}}(P)$ (identifié à un sous-groupe de $\mathfrak{S}_5$ après le choix d'une numérotation de ses racines) contient un 5-cycle.

De plus, on a

$$X^5 - X - \overline{1} = (X^2 + X + \overline{1})(X^3 + X^2 + \overline{1}) \in \mathbb{F}_2[X].$$

Le groupe $\mathrm{Gal}_{\mathbb{Q}}(P)$ contient donc une permutation de type $(2, 3)$. Le cube de cette permutation est donc une transposition. Finalement, $\mathrm{Gal}_{\mathbb{Q}}(P)$ s'identifie à un sous-groupe de $\mathfrak{S}_5$ contenant un 5-cycle et une transposition. D'après l'exercice VIII-21, on déduit donc

$$\mathrm{Gal}_{\mathbb{Q}}(P) \simeq \mathfrak{S}_5.$$

Le reste de ce chapitre est consacré à l'étude de deux familles d'extensions galoisiennes.

4. Extensions cyclotomiques

Soit $n \geqslant 1$ un entier et soit K un corps. On suppose que $\mathrm{car}(K) \nmid n$. Le polynôme $P = X^n - 1 \in K[X]$ est alors séparable.

En effet, on a $P' = nX^{n-1}$. L'hypothèse sur K montre que $P' \neq 0$, et donc que X est le seul facteur irréductible unitaire de P'. Or, $X \nmid P$. Ainsi, P et P' sont premiers entre eux et P est séparable par le lemme XXVII-3.2.

1. On peut aussi lister les polynômes irréductibles unitaires de degré 2 et 3 de $\mathbb{F}_5[X]$ et vérifier qu'aucun d'entre eux ne divise P.

L'ensemble
$$\mu_n(\overline{K}) = \{x \in \overline{K} \mid x^n = 1\}$$
est donc un sous-groupe de $\overline{K}^\times$ d'ordre n. Il est alors cyclique par le théorème XII-2.8.

4.1. Définition. Soit $n \geqslant 1$ un entier et soit K un corps. On suppose que l'on a $\mathrm{car}(K) \nmid n$.

Un élément de $\mu_n(\overline{K})$ est appelé une *racine n-ième de l'unité.*

On dit que $\zeta \in \overline{K}$ est une *racine primitive n-ième de l'unité* si c'est un générateur de $\mu_n(\overline{K})$, i.e. si c'est un élément de $\overline{K}^\times$ d'ordre n.

Dans ce cas, si $\zeta \in \overline{K}$ est une racine primitive n-ième de l'unité, alors les racines n-ièmes de l'unité sont les éléments [2]
$$\zeta^m, \quad m \in \mathbb{Z},\ \mathrm{pgcd}(m,n) = 1.$$

4.2. Exemple. Si $K = \mathbb{Q}$, alors $\zeta_n = e^{\frac{2i\pi}{n}}$ est une racine primitive n-ième de l'unité.

4.3. Remarque. Si $\zeta \in \overline{K}$ est une racine primitive n-ième de l'unité, on a
$$X^n - 1 = \prod_{k=0}^{n-1} (X - \zeta^k),$$
puisque ζ engendre $\mu_n(\overline{K})$, qui est exactement l'ensemble des racines du polynôme $X^n - 1$ dans $\overline{K}$.

On va maintenant étudier les extensions engendrées par une racine primitive n-ième de l'unité, notée ζ_n dans la suite.

D'après ce qui précède, $K(\zeta_n)/K$ est le corps des racines de $X^n - 1$, qui est un polynôme séparable. L'extension $K(\zeta_n)/K$ est donc galoisienne par la remarque 3.2. En particulier, elle ne dépend pas de la racine primitive choisie (puisqu'en fait elle les contient toutes).

4.4. Théorème. *Soit $n \geqslant 1$ un entier. Soit K un corps tel que $\mathrm{car}(K) \nmid n$, et soit $\zeta_n \in \overline{K}$ une racine primitive n-ième de l'unité. Alors, $K(\zeta_n)/K$ est galoisienne, et $\mathrm{Gal}\big(K(\zeta_n)/K\big)$ s'identifie à un sous-groupe de $(\mathbb{Z}/n\mathbb{Z})^\times$.*

En particulier, $\mathrm{Gal}\big(K(\zeta_n)/K\big)$ est abélien, d'ordre divisant $\varphi(n)$.

Démonstration. On vient de constater que $K(\zeta_n)/K$ est galoisienne. De plus, pour tout $\sigma \in \mathrm{Gal}\big(K(\zeta_n)/K\big)$, $\sigma(\zeta_n)$ est un élément d'ordre n de $\overline{K}^\times$, car σ est un automorphisme.

2. Bien entendu, on peut se restreindre à $m \in [\![0, n-1]\!]$.

Il existe donc un entier $m_\sigma \in \mathbb{Z}$ premier à n tel que $\sigma(\zeta_n) = \zeta_n^{m_\sigma}$. Cet entier m_σ est unique modulo n, et la classe $\overline{m}_\sigma$ est donc un élément du groupe multiplicatif $(\mathbb{Z}/n\mathbb{Z})^\times$ bien défini, que l'on note $\rho(\sigma)$. Montrons que l'application

$$\begin{array}{rccc} \rho : & \mathrm{Gal}\big(K(\zeta_n)/K\big) & \longrightarrow & (\mathbb{Z}/n\mathbb{Z})^\times \\ & \sigma & \longmapsto & \overline{m}_\sigma \end{array}$$

est un morphisme de groupes injectif.

Soient $\sigma_1, \sigma_2 \in \mathrm{Gal}\big(K(\zeta_n)/K\big)$, et soient $m_1, m_2 \in \mathbb{Z}$ premiers à n tels que

$$\sigma_i(\zeta_n) = \zeta_n^{m_i}, \quad i = 1, 2.$$

On a alors

$$(\sigma_1\sigma_2)(\zeta_n) = \sigma_1(\zeta_n^{m_2}) = \sigma_1(\zeta_n)^{m_2} = \zeta_n^{m_1 m_2}.$$

On en déduit donc

$$\rho(\sigma\tau) = \overline{m_1 m_2} = \overline{m}_1\overline{m}_2 = \rho(\sigma)\rho(\tau).$$

Ainsi, ρ est bien un morphisme de groupes. Soit alors $\sigma \in \mathrm{Gal}\big(K(\zeta_n)/K\big)$ tel que $\rho(\sigma) = \overline{1}$. Cela signifie donc que $\sigma(\zeta_n) = \zeta_n$. L'argument habituel montre alors que $\sigma = \mathrm{Id}_{K(\zeta_n)}$. Le dernier point est clair.

Cela achève la démonstration. □

Profitons-en pour introduire une nouvelle définition.

4.5. Définition. Soit K un corps. Une extension L/K est dite *abélienne* (resp. *cyclique*), si L/K est galoisienne, et si $\mathrm{Gal}(L/K)$ est un groupe abélien (resp. cyclique).

On dit que L/K est une extension *cyclotomique* s'il existe un entier $n \geqslant 1$ tel que $\mathrm{car}(K) \nmid n$ et $L = K(\zeta_n)$.

4.6. Exemple. D'après le théorème précédent, toute extension cyclotomique est abélienne.

Le lemme XXV-4.11 montre que l'on peut avoir $\mathrm{Gal}\big(K(\zeta_n)/K\big) \simeq (\mathbb{Z}/n\mathbb{Z})^\times$. En revanche, ce n'est pas toujours le cas, comme on peut le voir en prenant un corps K contenant ζ_n, auquel cas $\mathrm{Gal}\big(K(\zeta_n)/K\big)$ est trivial.

Il peut être assez délicat de déterminer $[K(\zeta_n) : K]$ et $\mathrm{Gal}\big(K(\zeta_n)/K\big)$ en général (le cas $K = \mathbb{Q}$ est déjà non trivial !). En revanche, lorsque K est un corps fini, on a une réponse complète, donnée par la proposition suivante.

4.7. Proposition. *Soit $q \geqslant 2$ une puissance d'un nombre premier. On suppose que q est premier à n. Soit d l'ordre de $\overline{q} \in (\mathbb{Z}/n\mathbb{Z})^\times$.*

Alors, $\mathbb{F}_q(\zeta_n) \simeq \mathbb{F}_{q^d}$, et le groupe $\mathrm{Gal}\big(\mathbb{F}_q(\zeta_n)/\mathbb{F}_q\big)$ est isomorphe au sous-groupe de $(\mathbb{Z}/n\mathbb{Z})^\times$ engendré par $\overline{q}$.

Démonstration. Puisque $\mathbb{F}_q(\zeta_n)$ est un corps fini, la proposition XXVI-2.11 montre que $\mathrm{Gal}\big(\mathbb{F}_q(\zeta_n)/\mathbb{F}_q\big)$ est cyclique, engendré par

$$\begin{array}{rccc} \sigma_q : & \mathbb{F}_q(\zeta_n) & \longrightarrow & \mathbb{F}_q(\zeta_n) \\ & x & \longmapsto & x^q. \end{array}$$

Si $\rho : \mathrm{Gal}\big(\mathbb{F}_q(\zeta_n)/\mathbb{F}_q\big) \longrightarrow (\mathbb{Z}/n\mathbb{Z})^\times$ est le morphisme injectif défini dans la démonstration du théorème précédent, on a donc $\rho(\sigma_q) = \overline{q}$. Ainsi, on obtient un isomorphisme

$$\mathrm{Gal}\big(\mathbb{F}_q(\zeta_n)/\mathbb{F}_q\big) \simeq \langle \overline{q} \rangle.$$

En particulier, on a

$$[\mathbb{F}_q(\zeta_n) : \mathbb{F}_q] = |\mathrm{Gal}\big(\mathbb{F}_q(\zeta_n)/\mathbb{F}_q\big)| = d.$$

Mais alors, $\mathbb{F}_q(\zeta_n)$ est un corps à q^d éléments, donc isomorphe à $\mathbb{F}_{q^d}$. □

Pour achever cette section, nous allons étudier le comportement des extensions $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ par composition et par intersection. Puisque l'extension cyclotomique $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ ne dépend pas du choix de la racine n-ième de l'unité, on peut supposer que $\zeta_n = e^{\frac{2i\pi}{n}}$, ce que l'on fera implicitement dans la suite.

4.8. Théorème. *Soient $m, n \geqslant 1$ deux entiers. Alors, on a*

$$\mathbb{Q}(\zeta_m)\,\mathbb{Q}(\zeta_n) = \mathbb{Q}(\zeta_D) \quad \text{et} \quad \mathbb{Q}(\zeta_m) \cap \mathbb{Q}(\zeta_n) = \mathbb{Q}(\zeta_d),$$

où $d = \mathrm{pgcd}(m, n)$ et $D = \mathrm{ppcm}(m, n)$.

Démonstration. Écrivons $D = rm = sn$, avec r et s premiers entre eux.

D'une part, on a $\zeta_m = \zeta_D^r$ et $\zeta_n = \zeta_D^s$. Ainsi, on a $\zeta_m, \zeta_n \in \mathbb{Q}(\zeta_D)$. Par conséquent, on obtient

$$\mathbb{Q}(\zeta_m, \zeta_n) = \mathbb{Q}(\zeta_m)\,\mathbb{Q}(\zeta_n) \subset \mathbb{Q}(\zeta_D).$$

D'autre part, il existe $u, v \in \mathbb{Z}$ tels que $ur + vs = 1$. On a alors

$$\zeta_D = (\zeta_D^r)^u (\zeta_D^s)^v = \zeta_m^u \zeta_n^v \in \mathbb{Q}(\zeta_m)\mathbb{Q}(\zeta_n).$$

Ainsi, on a $\mathbb{Q}(\zeta_D) \subset \mathbb{Q}(\zeta_m)\,\mathbb{Q}(\zeta_n)$, d'où la première égalité.

Posons $K = \mathbb{Q}(\zeta_m) \cap \mathbb{Q}(\zeta_n)$. Écrivons $m = kd$ et $n = \ell d$, avec k et ℓ premiers entre eux.

On a alors
$$\zeta_d = \zeta_m^k \in \mathbb{Q}(\zeta_m),$$
ainsi que
$$\zeta_d = \zeta_n^\ell \in \mathbb{Q}(\zeta_n).$$

On en déduit que $\mathbb{Q}(\zeta_d) \subset \mathbb{Q}(\zeta_m) \cap \mathbb{Q}(\zeta_n)$. Comme les extensions $\mathbb{Q}(\zeta_m)/\mathbb{Q}$ et $\mathbb{Q}(\zeta_n)/\mathbb{Q}$ sont galoisiennes, $\mathbb{Q}(\zeta_m)/K$ et $\mathbb{Q}(\zeta_n)/K$ sont linéairement disjointes par la proposition 2.11. On a alors

$$[\mathbb{Q}(\zeta_m)\mathbb{Q}(\zeta_n) : K] = [\mathbb{Q}(\zeta_D) : K] = [\mathbb{Q}(\zeta_m) : K][\mathbb{Q}(\zeta_n) : K].$$

En multipliant par $[K : \mathbb{Q}]^2$, et en utilisant la formule de multiplicativité des degrés, on obtient
$$\varphi(D)[K : \mathbb{Q}] = \varphi(m)\varphi(n).$$

Comme $[K : \mathbb{Q}] = [K : \mathbb{Q}(\zeta_d)] \times [\mathbb{Q}(\zeta_d) : \mathbb{Q}]$, on obtient finalement

$$[K : \mathbb{Q}(\zeta_d)] = \frac{\varphi(m)\varphi(n)}{\varphi(d)\varphi(D)}.$$

Il nous faut donc démontrer que cette dernière quantité est égale 1.

Soit T_1 l'ensemble des diviseurs premiers de m, et T_2 l'ensemble des diviseurs premiers de n. On a donc

$$\varphi(m) = m \prod_{p \in T_1} \left(1 - \frac{1}{p}\right) \quad \text{et} \quad \varphi(n) = n \prod_{p \in T_2} \left(1 - \frac{1}{p}\right),$$

ainsi que

$$\varphi(d) = d \prod_{p \in T_1 \cap T_2} \left(1 - \frac{1}{p}\right) \quad \text{et} \quad \varphi(D) = D \prod_{p \in T_1 \cup T_2} \left(1 - \frac{1}{p}\right).$$

On a ainsi

$$\varphi(d)\varphi(D) = dD\Big(\prod_{p \in T_1 \cap T_2} \left(1 - \frac{1}{p}\right)\Big)^2 \Big(\prod_{p \in T_1 \setminus T_2} \left(1 - \frac{1}{p}\right)\Big)\Big(\prod_{p \in T_2 \setminus T_1} \left(1 - \frac{1}{p}\right)\Big).$$

Mais, on a aussi

$$\varphi(m) = m\Big(\prod_{p \in T_1 \cap T_2} \left(1 - \frac{1}{p}\right)\Big)\Big(\prod_{p \in T_1 \setminus T_2} \left(1 - \frac{1}{p}\right)\Big),$$

ainsi que

$$\varphi(n) = n\Big(\prod_{p \in T_1 \cap T_2} \left(1 - \frac{1}{p}\right)\Big)\Big(\prod_{p \in T_2 \setminus T_1} \left(1 - \frac{1}{p}\right)\Big).$$

Par conséquent, on a

$$\varphi(m)\varphi(n) = mn\Big(\prod_{p \in T_1 \cap T_2} \left(1 - \frac{1}{p}\right)\Big)^2 \Big(\prod_{p \in T_1 \setminus T_2} \left(1 - \frac{1}{p}\right)\Big)\Big(\prod_{p \in T_2 \setminus T_1} \left(1 - \frac{1}{p}\right)\Big).$$

Comme $mn = dD$, on a l'égalité désirée, ce qui achève la démonstration. □

5. Théorie de Kummer

Nous allons nous intéresser dans cette section aux extensions cycliques.

5.1. Exemples

(1) Si $q \geqslant 2$ est une puissance d'un nombre premier, l'extension $\mathbb{F}_{q^d}/\mathbb{F}_q$ est cyclique pour tout $d \geqslant 1$ (cf. exemple 2.3 (2)).

(2) Pour tout nombre premier p impair, et tout $r \geqslant 1$, $\mathbb{Q}(\zeta_{p^r})/\mathbb{Q}$ est cyclique, puisqu'elle est galoisienne, de groupe de Galois isomorphe au groupe $(\mathbb{Z}/p^r\mathbb{Z})^\times$, qui est cyclique par le théorème XVII-2.4.

(3) Soit K un corps et soit $n \geqslant 1$ un entier. On suppose que $\mathrm{car}(K) \nmid n$, et que $\mu_n(\overline{K}) \subset K$ (ou ce qui revient au même, que K contient une racine primitive n-ième de l'unité).

Soit $a \in K^\times$. Alors, le polynôme $P = X^n - a \in K[X]$ est séparable. En effet, si $\alpha \in \overline{K}$ est une racine de P, on a

$$P = \prod_{\zeta \in \mu_n(\overline{K})} (X - \zeta\alpha).$$

Puisque $\mu_n(\overline{K}) \subset K$, on remarque alors que l'extension de K engendrée par une racine de P ne dépend pas de la racine choisie. On note cette extension $K(\sqrt[n]{a})/K$.

Autrement dit, c'est l'extension engendrée par un élément $\alpha \in \overline{K}$ vérifiant $\alpha^n = a$. On écrira d'ailleurs souvent $\alpha = \sqrt[n]{a}$ par abus dans la suite.

Tout ce qui précède montre que $K(\sqrt[n]{a}) = \mathrm{Dec}_K(P)$. Comme P est séparable, $K(\sqrt[n]{a})/K$ est galoisienne par la remarque 3.2. De plus, si $\sigma \in \mathrm{Gal}\big(K(\sqrt[n]{a})/K\big)$, on a

$$\sigma(\sqrt[n]{a})^n = \sigma(\sqrt[n]{a}^n) = \sigma(a) = a.$$

Ainsi, $\sigma(\sqrt[n]{a})$ est une racine P, et il existe une unique racine n-ième de l'unité ζ telle que $\sigma(\sqrt[n]{a}) = \zeta\sqrt[n]{a}$. On a donc une application

$$\begin{aligned} \varphi : \mathrm{Gal}\big(K(\sqrt[n]{a})/K\big) &\longrightarrow \mu_n(\overline{K}) \\ \sigma &\longmapsto \frac{\sigma(\sqrt[n]{a})}{\sqrt[n]{a}}. \end{aligned}$$

De plus, pour tous $\sigma, \tau \in \mathrm{Gal}\big(K(\sqrt[n]{a})/K\big)$, on a

$$\begin{aligned}\varphi(\sigma\tau) &= \frac{(\sigma\tau)(\sqrt[n]{a})}{\sqrt[n]{a}} \\ &= \frac{\sigma(\varphi(\tau)\sqrt[n]{a})}{\sqrt[n]{a}} \\ &= \varphi(\tau)\frac{\sigma(\sqrt[n]{a})}{\sqrt[n]{a}} \\ &= \varphi(\sigma)\varphi(\tau),\end{aligned}$$

l'avant-dernière égalité provenant du fait que $\varphi(\tau) \in \mu_n(\overline{K}) \subset K$. Ainsi, φ est un morphisme de groupes. De plus, φ est injectif. En effet, si $\sigma \in \mathrm{Gal}\big(K(\sqrt[n]{a})/K\big)$ vérifie $\varphi(\sigma) = 1$, on a $\sigma(\sqrt[n]{a}) = \sqrt[n]{a}$, et par suite $\sigma = \mathrm{Id}$.

Ainsi, $\mathrm{Gal}\big(K(\sqrt[n]{a})/K\big)$ est isomorphe à un sous-groupe du groupe cyclique $\mu_n(\overline{K})$. Par conséquent, $\mathrm{Gal}\big(K(\sqrt[n]{a})/K\big)$ est cyclique d'ordre divisant n. En particulier, $[K(\sqrt[n]{a}) : K] \mid n$.

5.2. Définition. Une *extension de Kummer* est une extension de corps de la forme $K(\sqrt[n]{a})/K$, où $\mathrm{car}(K) \nmid n$ et $\mu_n(\overline{K}) \subset K$.

5.3. Exemple. Soit K un corps de caractéristique différente de 2. Alors, toute extension L/K de degré $\leqslant 2$ est une extension de Kummer. Autrement dit, si $[L : K] \leqslant 2$, on a $L = K(\sqrt{d})$, avec $d \in K^\times$.

Ce fait a été déjà établi plus ou moins explicitement lorsque $K = \mathbb{Q}$, mais il est suffisamment essentiel dans les applications qu'il est bon de le souligner encore une fois.

Si $[L : K] = 1$, on a $L = K$ et on peut prendre $d = 1$.

Supposons que $[L : K] = 2$. Alors, $L \neq K$, et l'on peut choisir $\alpha \in L \setminus K$. Comme $\alpha \notin K$, $\mu_{\alpha,K}$ est de degré $\geqslant 2$. On a alors

$$K \subset K(\alpha) \subset L,$$

et puisque L et $K(\alpha)$ ont même degré sur K, on a $L = K(\alpha)$.

Écrivons $\mu_{\alpha,K} = X^2 + aX + b$, $a, b \in K$, et posons $d = a^2 - 4b$. Alors, on a

$$0 = \alpha^2 + a\alpha + b = \Big(\alpha + \frac{a}{2}\Big)^2 - \frac{d}{4}.$$

Ainsi, $\omega = 2\alpha + a \in L$ et vérifie $\omega^2 = d$. On a alors $K(\omega) \subset K(\alpha)$. Comme on a $\alpha = \dfrac{-a + \omega}{2}$, on obtient aussi $K(\alpha) \subset K(\omega)$, et par conséquent

$$L = K(\alpha) = K(\omega) = K(\sqrt{d}).$$

D'après l'exemple 5.1 (3), une extension de Kummer est cyclique. Nous allons démontrer que la réciproque est vraie. Nous aurons besoin d'un théorème dû à Hilbert.

5.4. Théorème (Théorème 90 de Hilbert). *Soit L/K une extension cyclique, et soit σ un générateur de $\mathrm{Gal}(L/K)$.*

Alors, pour tout $x \in L$, on a $N_{L/K}(x) = 1$ si, et seulement si, il existe un élément $y \in L^\times$ tel que $x = \dfrac{y}{\sigma(y)}$.

Démonstration. Soit $n = [L : K]$. Supposons tout d'abord qu'il existe un élément $y \in L^\times$ tel que $x = \dfrac{y}{\sigma(y)}$. Comme $\mathscr{X}(L/K) = \mathrm{Gal}(L/K) = \langle\sigma\rangle$, d'après la proposition XXVII-4.5, on a

$$N_{L/K}(x) = \prod_{k=0}^{n-1} \sigma^k(x) = \prod_{k=0}^{n-1} \frac{\sigma^k(y)}{\sigma^{k+1}(y)} = \frac{y}{\sigma^n(y)}.$$

Puisque $\sigma^n = \mathrm{Id}_L$, on obtient $N_{L/K}(x) = 1$.

Inversement, si $N_{L/K}(x) = 1$, on a $x, \sigma(x), \dots, \sigma^{n-1}(x) \in L^\times$. D'après le corollaire XXVII-1.5, la famille $(\mathrm{Id}_L, \sigma, \dots, \sigma^{n-1})$ est L-libre. En particulier, on a

$$\mathrm{Id}_L + x\sigma + \big(x\sigma(x)\big)\sigma^2 + \cdots + \big(x\sigma(x)\cdots\sigma^{n-2}(x)\big)\sigma^{n-1} \neq 0.$$

Il existe donc $z \in L$ tel que

$$y = z + x\sigma(z) + \big(x\sigma(x)\big)\sigma^2(z) + \cdots + \big(x\sigma(x)\dots\sigma^{n-2}(x)\big)\sigma^{n-1}(z) \neq 0.$$

Mais alors, en tenant compte du fait que $\sigma^n = \mathrm{Id}_L$, on obtient

$$\sigma(y) = \sigma(z) + \cdots + \big(\sigma(x)\cdots\sigma^{n-2}(x)\big)\sigma^{n-1}(z) + \big(\sigma(x)\cdots\sigma^{n-1}(x)\big)z.$$

Puisque $x\sigma(x)\cdots\sigma^{n-1}(x) = N_{L/K}(x) = 1$, on a finalement $x\sigma(y) = y$, soit encore $x = \dfrac{y}{\sigma(y)}$. Cela achève la démonstration. □

Nous pouvons maintenant démontrer le résultat principal de cette section, et qui généralise une partie des résultats de l'exercice XXVI-4.

5.5. Théorème (Kummer). *Soit $n \geqslant 1$ un entier. Soit K un corps tel que $\mathrm{car}(K) \nmid n$ et $\mu_n(\overline{K}) \subset K$. Alors, une extension L/K est cyclique de degré divisant n si, et seulement si, il existe $a \in K^\times$ tel que $L = K(\sqrt[n]{a})$.*

De plus, pour tous $a, b \in K^\times$, on a $K(\sqrt[n]{a}) = K(\sqrt[n]{b})$ si, et seulement si, il existe $m \in \mathbb{Z}$ premier à n et $u \in K^\times$ tels que $b = a^m u^n$.

Démonstration. L'exemple 5.1 (3) montre que $K(\sqrt[n]{a})/K$ est cyclique de degré divisant n. Réciproquement, soit L/K une extension cyclique de degré $d \mid n$. Par hypothèse, K contient les racines n-ièmes de l'unité. Soit $\zeta_d \in \mu_n(\overline{K}) \subset K$ un élément d'ordre d. On a alors

$$N_{L/K}(\zeta_d^{-1}) = \zeta_d^{-d} = 1.$$

Par le théorème 90 de Hilbert, il existe $\alpha \in L^\times$ tel que $\zeta_d^{-1} = \dfrac{\alpha}{\sigma(\alpha)}$, où σ est un générateur de $\mathrm{Gal}(L/K)$. Comme $\zeta_d \in \mu_n(\overline{K})$, on a

$$\zeta_d^{-n} = 1 = \frac{\alpha^n}{\sigma(\alpha^n)},$$

soit $\sigma(\alpha^n) = \alpha^n$. Puisque σ engendre $\mathrm{Gal}(L/K)$, on en déduit que α^n est fixe par tout élément du groupe $\mathrm{Gal}(L/K)$. Par le lemme 2.10, on a donc $\alpha^n \in K^\times$.

Posons $a = \alpha^n$, et vérifions que $L = K(\alpha) = K(\sqrt[n]{a})$. Bien entendu, on a l'inclusion $K(\sqrt[n]{a}) \subset L$. Par construction, on a $\sigma(\sqrt[n]{a}) = \zeta_d \sqrt[n]{a}$, et par récurrence

$$\sigma^i(\sqrt[n]{a}) = \zeta_d^i \sqrt[n]{a}, \quad \text{pour tout } i \in [\![0, d-1]\!].$$

Comme ζ_d est d'ordre d, ces éléments dont deux à deux distincts. Les restrictions de $\mathrm{Id}_L, \sigma, \ldots, \sigma^{d-1}$ fournissent donc d plongements de $K(\sqrt[n]{a})/K$ deux à deux distincts. On a alors

$$[K(\sqrt[n]{a}) : K] \geqslant |\mathscr{X}(K(\sqrt[n]{a})/K)| \geqslant d$$

par le théorème XXVII-1.11, par exemple. Mais alors, puisque $[L : K] = d$, on obtient $[K(\sqrt[n]{a}) : K] = [L : K]$, et ainsi $L = K(\sqrt[n]{a})$.

Supposons maintenant que $L = K(\sqrt[n]{a}) = K(\sqrt[n]{b})$. Soit σ un générateur du groupe $\mathrm{Gal}(L/K)$. D'après l'exemple 5.1 (3), il existe $\zeta \in \mu_n(\overline{K})$ tel que

$$\sigma(\sqrt[n]{a}) = \zeta \sqrt[n]{a}.$$

Comme σ est d'ordre d, des calculs similaires aux précédents montrent que ζ est d'ordre d. De même, il existe $\zeta' \in \mu_n(\overline{K})$ d'ordre d tel que

$$\sigma(\sqrt[n]{b}) = \zeta' \sqrt[n]{b}.$$

Puisque $\mu_d(\overline{K})$ est cyclique[3] d'ordre d, il existe donc $r \in \mathbb{Z}$ premier à d tel que $\zeta' = \zeta^r$. Nous allons montrer qu'il existe un entier $m \in \mathbb{Z}$ premier à n tel que $\zeta' = \zeta^m$.

Soit t le produit des nombres premiers p tels que $p \nmid d$ et $p \mid n$. Les entiers t et d étant premiers entre eux, le théorème chinois montre l'existence d'un entier $m \in \mathbb{Z}$ tel que $m \equiv r\ [d]$ et $m \equiv 1\ [t]$.

3. En effet, puisque $\mathrm{car}(K) \nmid n$, on a aussi $\mathrm{car}(K) \nmid d$.

Vérifions que m est premier à n. Soit p un diviseur premier de n. Si $p \nmid d$, alors $p \mid t$, et ainsi $p \nmid m$ car $m \equiv 1\ [t]$. Si maintenant $p \mid d$, on a $p \nmid r$ car r est premier à d. Mais alors, $p \nmid m$ puisque $m \equiv r\ [d]$.

Ainsi, m et n sont bien premiers entre eux. De plus, on a $m = r + ds$, $s \in \mathbb{Z}$. Mais alors, on a

$$\zeta' = \zeta^{m-ds} = \zeta^m.$$

Par conséquent, on obtient

$$\sigma\Big(\frac{\sqrt[n]{b}}{(\sqrt[n]{a})^m}\Big) = \frac{\zeta^m \sqrt[n]{b}}{\zeta^m (\sqrt[n]{a})^m} = \frac{\sqrt[n]{b}}{(\sqrt[n]{a})^m}.$$

Ainsi, $u = \dfrac{\sqrt[n]{b}}{(\sqrt[n]{a})^m}$ est fixé par σ, donc par tout élément de $\mathrm{Gal}(L/K)$. Par conséquent, $u \in K^\times$, et $b = a^m u^n$.

Inversement, si $b = a^m u^n$, avec $m \in \mathbb{Z}$ premier à n et $u \in K^\times$, on a

$$K(\sqrt[n]{b}) = K\big(u(\sqrt[n]{a})^m\big) = K\big((\sqrt[n]{a})^m\big) \subset K(\sqrt[n]{a}).$$

Puisque m est premier à n, il existe k, $\ell \in \mathbb{Z}$ tels que $km + \ell n = 1$. On a alors $a = (a^m)^k (a^\ell)^n$, et donc

$$u^{nk} a = b^k a^{\ell n},$$

d'où $a = b^k (a^\ell u^{-k})^n$. Comme précédemment, on obtient

$$K(\sqrt[n]{a}) \subset K(\sqrt[n]{b}),$$

d'où l'égalité désirée. Cela achève la démonstration. □

5.6. Remarque. Lorsque $n = 2$, la condition $\mu_2(\overline{K}) \subset K$ est automatiquement vérifiée. En tenant compte du fait qu'un entier premier à 2 est de la forme $2k + 1$, $k \in \mathbb{Z}$, on obtient le fait suivant. Pour tout corps K de caractéristique différente de 2, et tous a, $b \in K^\times$, on a

$$K(\sqrt{a}) = K(\sqrt{b}) \iff \text{il existe } u \in K^\times, \text{ tel que } b = au^2,$$

ce que l'on peut démontrer directement.

6. Exercices

E1. Soit L/K une extension de degré fini. On suppose connus les définitions et résultats de l'exercice XXVII-7.

a. Montrer que, pour tout $x \in L$, on a $\sigma(x) = x$ pour tout $\sigma \in \mathscr{X}(L/K)$ si, et seulement si, x est purement inséparable sur K.

b. Soit L/K une extension normale. Montrer que $M = L^{\mathrm{Gal}(L/K)}$ est l'ensemble des éléments de L purement inséparables sur K. Montrer que M/K est inséparable et que L/M est séparable.

E2. Soit $K = \mathbb{F}_2(T)$, et soit $L = K(\sqrt[3]{T}, \sqrt{T})$.

a. En considérant les sous-extensions $K(\sqrt[3]{T})/K$ et $K(\sqrt{T})/K$, montrer que $[L : K] = 6$.

b. Montrer que L/K n'est pas séparable.

c. Soit $Q = X^2 + X + 1 \in K[X]$.

(i) Vérifier que Q est irréductible.

(ii) Soit $j \in \overline{K}$ une racine de Q. Montrer que $j \notin K(\sqrt{T})$, et en déduire que $[K(\sqrt{T}, j) : K] = 4$.

d. Vérifier que les racines de $X^3 - T \in K[X]$ dans $\overline{K}$ sont

$$\sqrt[3]{T}, \quad j\sqrt[3]{T} \quad \text{et} \quad j^2\sqrt[3]{T}.$$

e. Déduire des questions précédentes que L/K n'est pas normale.

E3. Soient L_1/K et L_2/K deux sous-extensions de $\overline{K}/K$. On suppose que l'extension L_1/K est normale.

Montrer que $L_1 \simeq L_2$ si, et seulement si, $L_1 = L_2$. Est-ce encore vrai si l'on ne suppose plus que L_1/K est normale ?

E4. Soit $P \in K[X]$ un polynôme unitaire séparable. On écrit $P = P_1 \cdots P_r$, où $P_1, \ldots, P_r$ sont irréductibles unitaires deux à deux distincts. On note $L = \mathrm{Dec}_K(P)$ et $L_i = \mathrm{Dec}_K(P_i)$. On note aussi X_i l'ensemble des racines de P_i et X l'ensemble des racines de P dans $\overline{K}$.

a. Montrer que les applications

$$\begin{aligned} \mathrm{Gal}_K(P) &\longrightarrow \mathrm{Gal}_K(P_1) \times \cdots \times \mathrm{Gal}_K(P_r) \\ \sigma &\longmapsto (\sigma_{|L_1}, \ldots, \sigma_{|L_r}) \end{aligned}$$

et

$$\begin{aligned} \mathrm{Gal}_K(P) &\longrightarrow \mathfrak{S}(X_1) \times \cdots \times \mathfrak{S}(X_r) \\ \sigma &\longmapsto (\sigma_{|X_1}, \ldots, \sigma_{|X_r}) \end{aligned}$$

sont des morphismes de groupes injectifs.

b. Justifier qu'il existe un unique élément $s_i \in \mathfrak{S}(X)$ tel que

$$(s_i)_{|X_i} = \sigma_{|X_i} \quad \text{et} \quad (s_i)_{|X_j} = \mathrm{Id}_{|X_i}, \quad \text{si } i \neq j.$$

Montrer que $s_1, \ldots, s_r$ commutent deux à deux, et que

$$\sigma_{|X} = s_1 \cdots s_r \in \mathfrak{S}(X).$$

c. Montrer que $[\mathrm{Dec}_K(P) : K] \mid \deg(P_1)! \times \cdots \times \deg(P_r)!$.

E5. a. Soit $P \in \mathbb{Z}[X]$ un polynôme unitaire, et soit $\alpha \in \overline{K}$ une racine de P. On suppose que $\mathbb{Q}(\alpha)/\mathbb{Q}$ est galoisienne, et qu'il existe un nombre premier p tel que $\overline{P}$ soit irréductible. Montrer que $\mathrm{Gal}_{\mathbb{Q}}(P)$ est cyclique.

b. Soient $a, b \in \mathbb{Q}^\times$. On suppose que a, b, $ab \notin \mathbb{Q}^{\times 2}$. Soit $L = \mathbb{Q}(\sqrt{a}, \sqrt{b})$.

(i) Montrer que L/K est galoisienne de degré 4, et que

$$\mathrm{Gal}(L/\mathbb{Q}) \simeq \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}/2\mathbb{Z}.$$

(ii) Montrer que $\mathbb{Q}(\sqrt{a} + \sqrt{b}) = \mathbb{Q}(\sqrt{a}, \sqrt{b})$.

(iii) En déduire que $X^4 - 2(a+b)X^2 + (a-b)^2 \in \mathbb{Q}[X]$ est irréductible.

(iv) On suppose que $a, b \in \mathbb{Z}$. Utiliser la question **a.** pour retrouver le fait que $X^4 - 2(a+b)X^2 + (a-b)^2$ n'est pas irréductible modulo p.

E6. Soit L/K une extension de degré fini.

a. On suppose que L/K est séparable. On rappelle que la forme bilinéaire

$$\begin{aligned} b_L : L \times L &\longrightarrow K \\ (x, y) &\longmapsto \mathrm{Tr}_{L/K}(xy) \end{aligned}$$

est non dégénérée.

(i) En utilisant le lemme XXVII-4.8, montrer que pour toute K-base $\mathbf{e}$ de L, on a $\det\big(\mathrm{Mat}(b_L, \mathbf{e})\big) \in K^\times$.

(ii) Montrer alors que la classe de $\det\big(\mathrm{Mat}(b_L, \mathbf{e})\big)$ dans $K^\times/K^{\times 2}$ est indépendante du choix de $\mathbf{e}$.

On note

$$d_{L/K} = \overline{\det\big(\mathrm{Mat}(b_L, \mathbf{e})\big)} \in K^\times/K^{\times 2},$$

où $\mathbf{e}$ est une K-base arbitraire de L.

On notera $K\big(\sqrt{d_{L/K}}\big)/K$ l'extension correspondante.

On suppose jusqu'à la fin de l'exercice que L/K est galoisienne.

b. Montrer que $K\big(\sqrt{d_{L/K}}\big)/K$ est une sous-extension de L/K.

Indication. On pourra considérer un élément primitif de L/K.

Soit $\alpha \in L$ un élément primitif de L/K, et soit $P = \mu_{\alpha,K} \in K[X]$. On note X_P l'ensemble des racines de P dans $\overline{K}$.

c. Justifier que le morphisme de groupes

$$\begin{aligned} \mathrm{Gal}(L/K) &\longrightarrow \mathfrak{S}(X_P) \\ \sigma &\longmapsto \sigma_{|X_P} \end{aligned}$$

est injectif.

d. Vérifier que l'application

$$\begin{aligned} \mathrm{Gal}(L/K) &\longrightarrow X_P \\ \tau &\longmapsto \tau(\alpha) \end{aligned}$$

est bijective, et que pour tout $\sigma \in \mathrm{Gal}(L/K)$, les $\sigma_{|X_P}$-orbites de X sont en correspondance bijective avec les orbites de la permutation

$$\begin{aligned} \mathrm{Gal}(L/K) &\longrightarrow \mathrm{Gal}(L/K) \\ \tau &\longmapsto \sigma\tau. \end{aligned}$$

e. En utilisant l'exercice VIII-26 et le théorème 3.8, montrer que les propriétés suivantes sont équivalentes :

(1) on a $d_{L/K} = \overline{1} \in K^\times/K^{\times 2}$;

(2) l'image du morphisme de groupes injectif

$$\begin{aligned} \mathrm{Gal}(L/K) &\longrightarrow \mathfrak{S}(X_P) \\ \sigma &\longmapsto \sigma_{|X_P} \end{aligned}$$

est contenue dans $\mathfrak{A}(X_P)$;

(3) le groupe $\mathrm{Gal}(L/K)$ est un groupe d'ordre impair, ou est un groupe d'ordre pair dont un 2-Sylow n'est pas cyclique.

E7. Soit L/K une extension galoisienne. On se propose de démontrer le théorème de la base normale, à savoir qu'il existe $x \in L$ tel que la famille $\big(\sigma(x)\big)_{\sigma\in\mathrm{Gal}(L/K)}$ soit une K-base de L.

On suppose ici que K est infini, le cas des corps finis étant l'objet de l'exercice XXXI-12. Soit $G = \mathrm{Gal}(L/K)$.

a. Soit $(e_\sigma)_{\sigma\in G}$ une K-base de L. On ordonne les éléments de G, de façon à ce que le premier élément soit le morphisme identité. Montrer que la matrice

$$M = \big(\sigma(e_\tau)\big)_{\sigma,\tau}$$

est inversible, et en déduire l'existence d'éléments $\lambda_\rho \in L$, $\rho \in G$, tels que $\sum_\rho \lambda_\rho \sigma(e_\rho) = 1$ si $\sigma = \mathrm{Id}_L$ et 0 sinon.

b. Soit E un corps infini, soit $n \geqslant 1$ et soit $P \in E[T_1, \ldots, T_n]$. On suppose qu'il existe des sous-ensembles infinis $A_1, \ldots, A_n$ de E tels que P s'annule sur $A_1 \times \cdots \times A_n$. Montrer que $P = 0$.

c. Soit $P = \det\big(\sum_\rho X_\rho \tau^{-1}\sigma(b_\rho)\big)_{\sigma,\tau} \in L[X_\rho,\ \rho \in G]$.

(i) Montrer que $P\big((\lambda_\rho)_\rho\big) = 1$.

(ii) En déduire qu'il existe $(x_\rho)_\rho \in K^n$ tel que $P\big((x_\rho)_\rho\big) \neq 0$.

d. Montrer alors que la matrice $A = \big(\sum_\rho X_\rho \tau^{-1}\sigma(e_\rho)\big)_{\sigma,\tau}$ est inversible.

e. Soit $x = \sum_\rho x_\rho e_\rho$. Montrer que la famille $\big(\sigma(x)\big)_{\sigma\in G}$ est K-libre, et conclure.

E8. Soient m, $n \geqslant 1$ deux entiers, avec $m \geqslant n$. Montrer que $\mathbb{Q}(\zeta_n) = \mathbb{Q}(\zeta_m)$ si, et seulement si, $m = n$ ou n est impair et $m = 2n$.

Indication. Pour le sens non trivial, montrer que l'hypothèse implique l'égalité $\mathbb{Q}(\zeta_d) = \mathbb{Q}(\zeta_D)$, avec $d = \mathrm{pgcd}(m,n)$ et $D = \mathrm{ppcm}(m,n)$, puis que $\varphi(d) = \varphi(D)$.

E9. Soit p un nombre premier impair. On note $\zeta_p = e^{\frac{2i\pi}{p}}$.

a. Soit $s = \sum_{k=1}^{p-1} \left(\frac{k}{p}\right)\zeta_p^k$. Montrer que

$$s^2 = \sum_{\ell=1}^{p-1}\left(\frac{\ell}{p}\right)\Big(\sum_{k=1}^{p-1}\zeta_p^{k(1+\ell)}\Big).$$

En déduire que $s^2 = \left(\frac{-1}{p}\right)p = (-1)^{\frac{p-1}{2}}p$.

b. En déduire que $\mathbb{Q}(\sqrt{p^*}) \subset \mathbb{Q}(\zeta_p)$, où $p^* = (-1)^{\frac{p-1}{2}}p$.

c. Montrer que pour tout entier $d \in \mathbb{Z}\setminus\{0\}$, il existe un entier $n \geqslant 1$ tel que $\mathbb{Q}(\sqrt{d}) \subset \mathbb{Q}(\zeta_n)$.

E10. Soit L/K une extension cyclique, et soit σ un générateur de $\mathrm{Gal}(L/K)$. Montrer que pour tout $x \in L$, on a $\mathrm{Tr}_{L/K}(x) = 0$ si, et seulement si, il existe $y \in L$ tel que $x = y - \sigma(y)$.

Indication. En utilisant le lemme d'indépendance de Dedekind ou la séparabilité de L/K, montrer l'existence de $u \in L$ tel que

$$\mathrm{Tr}_{L/K}(u) = 1.$$

Considérer alors

$$y = \sum_{k=0}^{n-2}\Big(\sum_{\ell=0}^{k}\sigma^\ell(x)\Big)\sigma^k(u).$$

E11. Soit K un corps de caractéristique $p > 0$.

a. Soit $a \in K$. Soit $P = X^p - X - a \in K[X]$. Soit $\alpha \in \overline{K}$ une racine de P. Montrer que les racines de P dans $\overline{K}$ sont les éléments

$$\alpha + m\cdot 1_K, \quad m \in [\![0, p-1]\!].$$

Il s'ensuit que l'extension de corps $K(\alpha)/K$ est indépendante du choix d'une racine $\alpha \in \overline{K}$ de P. On la note $K\big(\wp^{-1}(a)\big)/K$.

b. On suppose que $P = X^p - X - a$ est irréductible[4]. Montrer que l'extension $K\big(\wp^{-1}(a)\big)/K$ est cyclique de degré p, et que $\mathrm{Gal}(L/K)$ est engendré par

$$\begin{aligned} L &\longrightarrow L \\ \alpha &\longmapsto \alpha + 1, \end{aligned}$$

où $\alpha \in \overline{K}$ est une racine de P fixée.

c. Soit L/K une extension cyclique de degré p, et soit σ un générateur de $\mathrm{Gal}(L/K)$.

d. En utilisant l'exercice précédent, montrer qu'il existe un élément $\alpha \in L$ tel que $\sigma(\alpha) = \alpha + 1_K$.

e. Justifier que $\alpha \notin K$, puisque $L = K(\alpha)$.

f. Calculer $\sigma(\alpha^p - \alpha)$, et en déduire que $a = \alpha^p - \alpha \in K$.

g. Montrer alors que $L = K\big(\wp^{-1}(a)\big)$.

h. Soient $a, b \in K$. Montrer que $K\big(\wp^{-1}(a)\big) = K\big(\wp^{-1}(b)\big)$ si, et seulement si, il existe $m \in \mathbb{Z}$ premier à p et $\lambda \in K$ tels que $b = ma + \lambda^p - \lambda$.

Indication. Soient α, $\beta \in \overline{K}$ tels que $\alpha^p - \alpha - a = 0$ et $\beta^p - \beta - b = 0$ et soit σ un générateur de $\mathrm{Gal}(L/K)$. Justifier qu'il existe k, $\ell \in \mathbb{Z}$ premiers à p tels que

$$\sigma(\alpha) = \alpha + k \quad \text{et} \quad \sigma(\beta) = \beta + \ell.$$

En déduire l'existence de $m \in \mathbb{Z}$ premier à p tel que $\lambda = \beta - m\alpha \in K$. Calculer $\lambda^p - \lambda$ et conclure.

E12. Soit $d \in \mathbb{Z} \setminus \{0\}$ sans facteurs carrés. En utilisant le théorème 90 de Hilbert, résoudre l'équation

$$x^2 - dy^2 = 1, \quad x, y \in \mathbb{Q}.$$

Le but des cinq exercices suivants est de démontrer que tout groupe fini d'ordre 8 peut être obtenu comme le groupe de Galois d'une extension galoisienne de $\mathbb{Q}$.

E13. Soit $\alpha = \sqrt{\dfrac{5+\sqrt{5}}{2}}$, soit $L = \mathbb{Q}(\alpha)$, et soit $E = \mathbb{Q}(\alpha, i)$.

a. Vérifier que $\mathbb{Q}(\sqrt{5}) \subset L$, et que $\dfrac{5+\sqrt{5}}{2}$ n'est pas un carré dans le corps $\mathbb{Q}(\sqrt{5})$. En déduire $[L : \mathbb{Q}]$.

Indication. On pourra utiliser la norme de $\mathbb{Q}(\sqrt{5})/\mathbb{Q}$.

4. Cela équivaut au fait que a n'est pas de la forme $\lambda^p - \lambda$, $\lambda \in K$, cela d'après l'exercice XXVI-7.

b. Montrer que $L/\mathbb{Q}$ est galoisienne, et que $\mathrm{Gal}(L/\mathbb{Q})$ est un groupe cyclique d'ordre 4, engendré par

$$\sigma : \left| \begin{array}{ccc} \sqrt{2} & \rightsquigarrow & -\sqrt{2} \\ \alpha & \rightsquigarrow & \beta, \end{array} \right.$$

où $\beta = \dfrac{5-\sqrt{5}}{2}$. On pourra remarquer que $\alpha\beta = \sqrt{5}$.

c. Quel est le groupe de Galois de $E/\mathbb{Q}$, à isomorphisme près ?

Indication. Que vaut $L \cap \mathbb{Q}(i)$?

E14. Montrer que $\mathbb{Q}(\sqrt{2}, \sqrt{3}, i)/\mathbb{Q}$ est galoisienne, de groupe de Galois isomorphe à $(\mathbb{Z}/2\mathbb{Z})^3$.

E15. Soit $P = X^4 - 2X^2 - 1 \in \mathbb{Q}[X]$, et soit $L = \mathrm{Dec}_{\mathbb{Q}}(P)$.

a. Vérifier que le polynôme P est irréductible, et que les racines de P dans $\mathbb{C}$ sont $\pm\alpha$, $\pm\beta$, avec

$$\alpha = \sqrt{1+\sqrt{2}} \quad \text{et} \quad \beta = \sqrt{1-\sqrt{2}}.$$

b. Montrer que $L = \mathbb{Q}(\alpha, i)$, et en déduire la valeur de $[L : \mathbb{Q}]$.

c. Montrer que $\mathbb{Q}(\sqrt{2}) \subset \mathbb{Q}(\alpha)$ et déterminer $\mu_{\alpha,\mathbb{Q}(\sqrt{2})}$.

d. En considérant la tour d'extensions

$$\mathbb{Q} \subset \mathbb{Q}(\sqrt{2}) \subset \mathbb{Q}(\sqrt{2}, \alpha) \subset \mathbb{Q}(\sqrt{2}, \alpha, i),$$

calculer les éléments de $\mathrm{Gal}(L/\mathbb{Q})$.

e. Montrer que $\mathrm{Gal}(L/\mathbb{Q}) \simeq D_4$.

E16. Soit $E = \mathbb{Q}(\zeta_{17})$ et $L = \mathbb{Q}(\zeta_{17} + \zeta_{17}^{-1})$. On rappelle que $E/\mathbb{Q}$ est galoisienne, et que $\mathrm{Gal}(E/\mathbb{Q})$ d'ordre 16, engendré par

$$\sigma : \zeta_{17} \rightsquigarrow \zeta_{17}^3.$$

a. Vérifier que $[E : L] = 2$.

b. Montrer que σ se restreint en un automorphisme de $L/\mathbb{Q}$. Calculer alors $\sigma^4(\zeta_{17} + \zeta_{17}^{-1})$, et en déduire que le groupe $\mathrm{Gal}(L/\mathbb{Q})$ contient un élément d'ordre 8.

c. Déduire des questions précédentes que $L/\mathbb{Q}$ est galoisienne, et que l'on a $\mathrm{Gal}(L/\mathbb{Q}) \simeq \mathbb{Z}/8\mathbb{Z}$.

E17. Soit $\alpha = (2+\sqrt{2})(3+\sqrt{6})$, et soit $K = \mathbb{Q}(\alpha)$.

a. Montrer que $K = \mathbb{Q}(\sqrt{2}, \sqrt{3})$. En déduire que $\mathrm{Gal}(K/\mathbb{Q})$ est galoisienne, de groupe de Galois isomorphe à $(\mathbb{Z}/2\mathbb{Z})^2$.

On note

$$\sigma : \left| \begin{array}{ccc} \sqrt{2} & \rightsquigarrow & \sqrt{2} \\ \sqrt{3} & \rightsquigarrow & -\sqrt{3} \end{array} \right. , \quad \tau : \left| \begin{array}{ccc} \sqrt{2} & \rightsquigarrow & -\sqrt{2} \\ \sqrt{3} & \rightsquigarrow & \sqrt{3} \end{array} \right. .$$

b. Montrer que $\dfrac{\sigma(\alpha)}{\alpha}$, $\dfrac{\tau(\alpha)}{\alpha}$ et $\dfrac{\sigma\tau(\alpha)}{\alpha}$ sont les carrés d'éléments de $K^\times$, que l'on précisera.

Soit $\theta = \sqrt{\alpha}$, et soit $L = \mathbb{Q}(\theta)$.

c. Montrer que $\theta \notin K$.

d. Montrer que $L/\mathbb{Q}$ est galoisienne de degré 8.

e. Expliciter un prolongement $\widetilde{\sigma} \in \mathrm{Gal}(L/\mathbb{Q})$ de σ à L. De même, expliciter un prolongement $\widetilde{\tau} \in \mathrm{Gal}(L/\mathbb{Q})$ de τ à L.

f. Montrer que $\widetilde{\sigma}$ et $\widetilde{\tau}$ sont d'ordre 4, et que $\widetilde{\sigma}\widetilde{\tau} = -\widetilde{\tau}\widetilde{\sigma}$.

g. En déduire que $\mathrm{Gal}(L/\mathbb{Q}) \simeq Q_8$.

E18. Soit p un nombre premier, et soit $P \in \mathbb{Q}[X]$ un polynôme irréductible possédant deux racines complexes conjuguées et $p-2$ racines réelles Si X est l'ensemble des racines de P dans $\mathbb{C}$, on identifie $\mathrm{Gal}_{\mathbb{Q}}(P)$ à un sous-groupe de $\mathfrak{S}(X)$.

a. En considérant la restriction de la conjugaison complexe à $\mathrm{Dec}_{\mathbb{Q}}(P)$, montrer que $\mathrm{Gal}_{\mathbb{Q}}(P)$ contient une transposition.

b. Montrer que $\mathrm{Gal}_{\mathbb{Q}}(P)$ contient un p-cycle.

Indication. Utiliser l'exercice VIII-34.

c. En déduire que $\mathrm{Gal}_{\mathbb{Q}}(P) \simeq \mathfrak{S}_p$.

E19. Soit $n \geqslant 2$ un entier.

a. Exhiber un polynôme irréductible unitaire de $\mathbb{Q}[X]$ dont le groupe de Galois est isomorphe à $\mathfrak{S}_2$, resp. à $\mathfrak{S}_3$.

b. Soit G un sous-groupe transitif de $\mathfrak{S}_n$ contenant un $(n-1)$-cycle et une transposition. Montrer que $G = \mathfrak{S}_n$.

Indication. Montrer que l'on peut supposer que G contient $\sigma = (2 \cdots n)$. Utiliser alors la transitivité de G pour montrer qu'il contient une transposition de la forme $\tau = (1\ i)$, où $i \in [\![2, n]\!]$. Calculer alors $\sigma^k \tau \sigma^{-k}$ pour tout $k \in \mathbb{Z}$.

On suppose dorénavant que $n \geqslant 4$.

c. Justifier l'existence de polynômes unitaires P_2, P_3, $P_5 \in \mathbb{Z}[X]$ de degré n tels que P_2 est irréductible modulo 2, P_3 a un facteur irréductible de degré $n-1$ modulo 3, et la réduction de P_5 modulo 5 a un facteur irréductible de degré 2 et tous ses autres facteurs irréductibles sont deux à deux distincts et de degrés impairs.

d. Soit $P = -15P_2 + 10P_3 + 6P_5$. Montrer que $P \in \mathbb{Z}[X]$ est unitaire de degré n, séparable, irréductible, et que $\mathrm{Gal}_{\mathbb{Q}}(P) \simeq \mathfrak{S}_n$.

E20. Soit $K = \mathbb{Q}(i\sqrt{3})$, et soit $P = X^2 - (2 + i\sqrt{3}) \in K[X]$. Soit $\theta \in \mathbb{C}$ une racine de P. On pose $L = K(\theta)$.

a. Montrer que P est irréductible sur K. En déduire $[L : \mathbb{Q}]$.

b. Montrer que $L/\mathbb{Q}$ est galoisienne si, et seulement si, $K(\theta) = K(\overline{\theta})$, où $\overline{\theta}$ est le conjugué complexe de θ.

c. En utilisant la théorie de Kummer, déterminer si $L/\mathbb{Q}$ est galoisienne ou non.

E21. On pose $K = \mathbb{Q}(i)$, $P = X^4 - (1-i)$ et $Q = X^4 - (17 + 31i)$. On désigne par $\alpha \in \mathbb{C}$ une racine de P, et par $\beta \in \mathbb{C}$ une racine de Q.

a. En considérant le carré du module, montrer que $1 - i$ et $17 + 31i$ ne sont pas des carrés dans K.

b. En utilisant la théorie de Kummer, en déduire que les deux extensions $K(\alpha)/K$ et $K(\beta)/K$ sont galoisiennes, de groupes de Galois isomorphes à $\mathbb{Z}/4\mathbb{Z}$.

c. En utilisant la théorie de Kummer, montrer que $K(\alpha) = K(\beta)$ si, et seulement si, une des deux conditions suivantes est vérifiée :

(i) il existe $u \in K^\times$ tel que $17 + 31i = (1-i)u^4$;

(ii) il existe $u \in K^\times$ tel que $17 + 31i = (1-i)^3 u^4$.

d. Montrer que (ii) est impossible.

e. Résoudre dans K les équations $z^2 = -7 + 24i$ et $z^2 = 3 + 4i$.
A-t-on $K(\alpha) = K(\beta)$?

Chapitre XXIX

Nul n'est censé ignorer Galois

Ce chapitre au titre empreint d'humour glacé et sophistiqué[1] vise à établir une correspondance bijective entre les sous-groupes de $\mathrm{Gal}(L/K)$ et les sous-extensions de L/K lorsque L/K est galoisienne (correspondance annoncée dans la section XXV-5), et d'en donner quelques applications classiques.

1. Théorie de Galois

Soit L/K une extension de corps de degré fini. On note $\mathscr{SE}(L/K)$ l'ensemble des sous-extensions de K, et $\mathscr{SG}\big(\mathrm{Gal}(L/K)\big)$ l'ensemble des sous-groupes de $\mathrm{Gal}(L/K)$. Le lemme XXV-4.10 (2) implique que l'on a une application

$$\begin{array}{rccc} \Phi_{L/K} : & \mathscr{SG}\big(\mathrm{Gal}(L/K)\big) & \longrightarrow & \mathscr{SE}(L/K) \\ & H & \longmapsto & L^H/K. \end{array}$$

Notons que l'on peut également définir une application en sens inverse. Si F/K est une sous-extension de L/K, le groupe $\mathrm{Gal}(L/F)$ est un sous-groupe de $\mathrm{Gal}(L/K)$, puisque si un automorphisme de L est F-linéaire, il est aussi K-linéaire. On a donc une application

$$\begin{array}{rccc} \Psi_{L/K} : & \mathscr{SE}(L/K) & \longrightarrow & \mathscr{SG}\big(\mathrm{Gal}(L/K)\big) \\ & F/K & \longmapsto & \mathrm{Gal}(L/F). \end{array}$$

1. Le fabuleux jeu de mots présent dans le titre m'a obligeamment été prêté par mon ex-étudiant Louis-Thibault Gauthier. Dépêchez-vous de rire, il faut que je le lui rende.

Il est naturel de se demander si ces applications sont bijectives, et inverses l'une de l'autre. Comme on l'a déjà constaté dans l'exemple XXV-5.1, l'application $\Phi_{L/K}$ n'est pas nécessairement surjective. Néanmoins, nous allons démontrer qu'elle est toujours injective.

Commençons sans plus attendre par le résultat-clef de cette section.

1.1. Théorème (Lemme d'Artin). *Soit L un corps, et soit H un sous-groupe fini du groupe* $\mathrm{Aut}(L)$ *des automorphismes de l'anneau L. Alors, l'extension L/L^H est galoisienne, et l'on a*

$$\mathrm{Gal}(L/L^H) = H.$$

Démonstration. Nous allons d'abord établir le fait suivant.

Fait. Soit $\alpha \in L$, et soit $S = \{\sigma_1, \ldots, \sigma_r\}$ un sous-ensemble maximal de H tels que $\sigma_1(\alpha), \ldots, \sigma_r(\alpha) \in L$ soient deux à deux distincts [2]. Quitte à remplacer S par $S' = \{\sigma_1^{-1}\sigma_j \mid j \in [\![1, r]\!]\}$, on peut supposer sans perte de généralité que $\mathrm{Id}_L \in S$ (les $\sigma_1^{-1}\sigma_j(\alpha)$ sont bien deux à deux distincts par injectivité de σ_1^{-1}).

Alors, $Q_\alpha = \prod_{i=1}^{r} \big(X - \sigma_i(\alpha)\big) \in L^H[X]$, Q_α est séparable sur L^H, et l'on a de plus $\mu_{\alpha,L^H} \mid Q_\alpha$.

Gardons les notations précédentes. Pour tout $\sigma \in H$, on a alors

$$\sigma(Q_\alpha) = \prod_{i=1}^{r} \big(X - \sigma\big(\sigma_i(\alpha)\big)\big).$$

Soit $i \in [\![1, r]\!]$. Puisque H est un groupe, on a encore $\sigma\sigma_i \in H$. Par maximalité de S, il existe $k \in [\![1, r]\!]$ tel que $\sigma\big(\sigma_i(\alpha)\big) = \sigma_k(\alpha)$.

De plus, par injectivité de σ, pour tous $i, j \in [\![1, r]\!]$, on a

$$\sigma\big(\sigma_i(\alpha)\big) = \sigma\big(\sigma_j(\alpha)\big) \quad \Longrightarrow \quad i = j,$$

puisque les éléments $\sigma_1(\alpha), \ldots, \sigma_r(\alpha)$ sont deux à deux distincts. Autrement dit, σ induit une permutation de l'ensemble $\{\sigma_1(\alpha), \ldots, \sigma_r(\alpha)\}$. On en déduit alors que $\sigma(Q_\alpha) = Q_\alpha$ pour tout $\sigma \in H$, d'où $Q_\alpha \in L^H[X]$.

Le fait que Q_α soit séparable sur L^H découle alors de sa définition. Enfin, puisque $\mathrm{Id}_L \in S$, on a $Q_\alpha(\alpha) = 0$. Comme $Q_\alpha \in L^H[X]$, on obtient finalement $\mu_{\alpha,L^H} \mid Q_\alpha$.

Montrons maintenant le théorème. D'après le fait précédent, dont on garde les notations, pour tout $\alpha \in L$, $\mu_{\alpha,L^H} \mid Q_\alpha$. Comme Q_α est séparable sur L^H, il en est de même de μ_{α,L^H}, et α est donc séparable sur L^H. Ainsi,

2. Les éléments $\sigma_1(\alpha), \ldots, \sigma_r(\alpha)$ sont bien dans L, puisque σ_i est un automorphisme de L.

l'extension L/L^H est séparable. De plus, puisque $\mu_{\alpha,L^H} \mid Q_\alpha$, on a

$$[L^H(\alpha) : L^H] \leqslant \deg(Q_\alpha) \leqslant |H| \text{ pour tout } \alpha \in L.$$

Montrons que l'extension L/L^H est de degré fini $\leqslant |H|$. Supposons au contraire que $[L : L^H] \geqslant |H| + 1$ (où $[L : L^H]$ est possiblement infini). Alors, il existe $|H|+1$ éléments $\alpha_1, \ldots, \alpha_{|H|+1} \in L$ qui sont L^H-linéairement indépendants. Mais alors, l'extension $L^H(\alpha_1, \ldots, \alpha_{|H|+1})/L^H$ est séparable, de degré fini $\geqslant |H| + 1$ (puisque les α_i sont K-linéairement indépendants). Par le théorème de l'élément primitif, elle est engendrée par un unique élément de L. Elle est donc de degré $\leqslant |H|$ par ce qui précède, d'où une contradiction.

Or, tout élément de H fixe les éléments de L^H, et donc $H \subset \mathrm{Gal}(L/L^H)$. On a ainsi

$$|\mathrm{Gal}(L/L^H)| \geqslant |H|.$$

Mais on a aussi

$$|\mathrm{Gal}(L/L^H)| \leqslant [L : L^H] \leqslant |H|$$

d'après le corollaire XXVII-1.12 et le point précédent. On déduit de tout cela que $|\mathrm{Gal}(L/L^H)| = [L : L^H] = |H|$, ce qui démontre que L/L^H est galoisienne de degré $|H|$ par le lemme XXVIII-2.1, puis que $H = \mathrm{Gal}(L/L^H)$. Cela achève la démonstration. □

Comme corollaire immédiat du lemme d'Artin, on obtient une nouvelle caractérisation des extensions galoisiennes.

1.2. Corollaire. *Soit L/K une extension de degré fini. Alors, L/K est galoisienne si, et seulement si, $L^{\mathrm{Gal}(L/K)} = K$.*

Démonstration. Le sens direct est donné par le lemme XXVIII-2.10. Supposons maintenant que $L^{\mathrm{Gal}(L/K)} = K$. Par le lemme d'Artin, on a

$$[L : K] = [L : L^{\mathrm{Gal}(L/K)}] = |\mathrm{Gal}(L/K)|.$$

Mais alors, L/K est galoisienne par le lemme XXVIII-2.1. □

Le lemme d'Artin nous donne également le résultat suivant.

1.3. Proposition. *Soit L/K une extension de degré fini. Alors, on a*

$$\Psi_{L/K} \circ \Phi_{L/K} = \mathrm{Id}_{\mathscr{SG}(\mathrm{Gal}(L/K))}.$$

En particulier, l'application

$$\begin{array}{rccc} \Phi_{L/K} : & \mathscr{SG}\big(\mathrm{Gal}(L/K)\big) & \longrightarrow & \mathscr{SE}(L/K) \\ & H & \longmapsto & L^H/K \end{array}$$

est injective.

De plus, si $\Phi_{L/K}$ est bijective, alors son inverse est l'application

$$\begin{aligned}\Psi_{L/K} : \mathscr{SE}(L/K) &\longrightarrow \mathscr{SG}\big(\mathrm{Gal}(L/K)\big)\\ F/K &\longmapsto \mathrm{Gal}(L/F).\end{aligned}$$

Dans ce cas, L/K est nécessairement galoisienne.

Démonstration. L'égalité $\Psi_{L/K}\circ\Phi_{L/K} = \mathrm{Id}_{\mathscr{SG}(\mathrm{Gal}(L/K))}$ découle du lemme d'Artin. Tout le reste de la proposition est alors clair, sauf le dernier point. Supposons que $\Phi_{L/K}$ soit bijective. En particulier, il existe un sous-groupe H de $\mathrm{Gal}(L/K)$ tel que $L^H = K$. Mais alors, on a

$$K \subset L^{\mathrm{Gal}(L/K)} \subset L^H = K,$$

d'où $L^{\mathrm{Gal}(L/K)} = K$. Le corollaire 1.2 montre alors que L/K est galoisienne, d'où le résultat. □

Continuons par établir quelques propriétés des applications $\Phi_{L/K}$ et $\Psi_{L/K}$, valables pour toute extension L/K de degré fini.

1.4. Lemme. *Soit L/K une extension de degré fini. Pour toutes sous-extensions F_1/K, F_2/K et F/K de L/K, et pour tous sous-groupes H_1, H_2 et H de $\mathrm{Gal}(L/K)$, on a*

(1) $\mathrm{Gal}(L/F) = \{\sigma \in \mathrm{Gal}(L/K) \mid \sigma_{|F} = \mathrm{Id}_F\}$;

(2) $\mathrm{Gal}(L/F_1F_2) = \mathrm{Gal}(L/F_1) \cap \mathrm{Gal}(L/F_2)$;

(3) $L^{H_1} \cap L^{H_2} = L^{\langle H_1, H_2\rangle}$;

(4) *pour tout $\sigma \in \mathrm{Gal}(L/K)$, on a $\sigma(L^H) = L^{\sigma H \sigma^{-1}}$.*

Démonstration. Gardons les notations du lemme.

Si $\sigma \in \mathrm{Gal}(L/F)$ est F-linéaire, elle est K-linéaire et donc $\sigma \in \mathrm{Gal}(L/K)$. De plus, σ fixe les éléments de F, puisque pour tout $x \in F$, on a

$$\sigma(x) = \sigma(x1) = x\sigma(1) = x1 = x.$$

Inversement, si $\sigma \in \mathrm{Gal}(L/K)$ et fixe les éléments de F, σ est F-linéaire. En effet, c'est un morphisme de groupes additifs, et de plus, pour tout $x \in F$ et tout $x' \in L$, on a

$$\sigma(xx') = \sigma(x)\sigma(x') = x\sigma(x').$$

Cela démontre (1).

Montrons le point (2). Par (1), il suffit de montrer que $\sigma \in \mathrm{Gal}(L/K)$ fixe les éléments de F_1F_2 si, et seulement si, σ fixe les éléments de F_1 et de F_2. Si $\sigma \in \mathrm{Gal}(L/K)$ fixe les éléments de F_1F_2, il fixe aussi les éléments de F_1 et de F_2. Inversement, si $\sigma \in \mathrm{Gal}(L/K)$ fixe les éléments de F_1 et de F_2,

alors σ fixe les éléments de L de la forme

$$\sum_{i=1}^{n} x_i y_i, \quad n \geqslant 0,\ x_i \in F_1,\ y_i \in F_2.$$

Comme F_1/K est de degré fini, elle est algébrique. Le lemme XXV-3.20 montre alors que les éléments de F_1F_2 sont exactement de la forme précédente, d'où (2).

Le point (3) provient du fait qu'un élément $x \in L$ est fixé par tous les éléments de $\langle H_1, H_2\rangle$ si, et seulement si, il est fixé par les éléments de H_1 et par les éléments de H_2.

Il reste à démontrer (4). Pour tout $x \in L$, et tout $\tau \in H$, on a

$$\sigma\big(\tau(x)\big) = \sigma\tau\sigma^{-1}\big(\sigma(x)\big).$$

En particulier, pour tout $x \in L^H$, on a $\sigma(x) = \sigma\tau\sigma^{-1}\big(\sigma(x)\big)$. Ainsi, on obtient

$$\sigma(L^H) \subset L^{\sigma H\sigma^{-1}}.$$

Si maintenant $y \in L^{\sigma H\sigma^{-1}}$, soit $x \in L$ tel que $y = \sigma(x)$. Pour tout $\tau \in H$, on a alors

$$\sigma\big(\tau(x)\big) = \sigma\tau\sigma^{-1}(y) = y = \sigma(x),$$

et par injectivité de σ, on obtient $\tau(x) = x$. Par conséquent, on a $x \in L^H$, et ainsi $y \in \sigma(L^H)$. Cela achève la démonstration. □

Nous allons maintenant démontrer que si L/K est galoisienne, les applications $\Phi_{L/K}$ et $\Psi_{L/K}$ sont inverses l'une de l'autre.

1.5. Théorème (Correspondance de Galois). *Soit L/K une extension galoisienne de degré fini. Alors, on a une correspondance bijective décroissante entre l'ensemble des sous-groupes de $\mathrm{Gal}(L/K)$ et l'ensemble des sous-extensions de L/K. Cette correspondance est donnée par*

$$\begin{array}{ccc} \mathscr{SG}\big(\mathrm{Gal}(L/K)\big) & \longleftrightarrow & \mathscr{SE}(L/K) \\ H & \overset{\Phi_{L/K}}{\longmapsto} & L^H/K \\ \mathrm{Gal}(L/F) & \overset{\Psi_{L/K}}{\longleftarrow} & F/K. \end{array}$$

De plus, pour tout sous-groupe H de $\mathrm{Gal}(L/K)$, et toute sous-extension F/K de L/K, on a

$$[F : K] = [\mathrm{Gal}(L/K) : \mathrm{Gal}(L/F)] \quad \text{et} \quad [L^H : K] = [\mathrm{Gal}(L/K) : H].$$

Enfin, pour toutes sous-extensions F_1/K, F_2/K et F/K de L/K correspondant respectivement aux sous-groupes H_1, H_2 et H de $\mathrm{Gal}(L/K)$, on a les propriétés suivantes :

(1) *la sous-extension F_1F_2/K correspond au sous-groupe $H_1 \cap H_2$;*

(2) *la sous-extension* $F_1 \cap F_2/K$ *correspond au sous-groupe* $\langle H_1, H_2 \rangle$;
(3) *pour tout* $\sigma \in \mathrm{Gal}(L/K)$, $\sigma(F)/K$ *correspond au sous-groupe* $\sigma H \sigma^{-1}$;
(4) *l'extension* F/K *est galoisienne si, et seulement si,* H *est un sous-groupe distingué de* $\mathrm{Gal}(L/K)$. *Dans ce cas, on a un isomorphisme de groupes*

$$\mathrm{Gal}(F/K) \simeq \mathrm{Gal}(L/K)/H.$$

Démonstration. Montrons que $\Phi_{L/K}$ est bijective, d'inverse $\Psi_{L/K}$. D'après la proposition 1.3, on sait déjà que

$$\Psi_{L/K} \circ \Phi_{L/K} = \mathrm{Id}_{\mathscr{SG}(\mathrm{Gal}(L/K))}.$$

Si maintenant F/K est une sous-extension de L/K, l'extension L/F est galoisienne d'après la proposition XXVIII-2.5. Mais alors, on a

$$(\Phi_{L/K} \circ \Psi_{L/K})(F/K) = L^{\mathrm{Gal}(L/F)}/K = F/K,$$

la dernière égalité étant donnée par le corollaire 1.2. Autrement dit, on a bien $\Phi_{L/K} \circ \Psi_{L/K} = \mathrm{Id}_{\mathscr{SE}(L/K)}$, d'où la bijectivité de la correspondance.

De plus, si $H_1 \subset H_2$, on a clairement $L^{H_2} \subset L^{H_1}$. De même, si $F_1 \subset F_2$, on a $\mathrm{Gal}(L/F_2) \subset \mathrm{Gal}(L/F_1)$, puisqu'un automorphisme F_2-linéaire de L est aussi F_1-linéaire. Ainsi, la correspondance est décroissante.

Si F/K est une sous-extension de L/K, comme L/F est galoisienne, on a

$$[L : F] = |\mathrm{Gal}(L/F)|.$$

Pour la même raison, on a $[L : K] = |\mathrm{Gal}(L/K)|$. Par conséquent, on a

$$[F : K] = \frac{[L : K]}{[L : F]} = \frac{|\mathrm{Gal}(L/K)|}{|\mathrm{Gal}(L/F)|} = [\mathrm{Gal}(L/K) : \mathrm{Gal}(L/F)].$$

L'autre égalité en découle par bijectivité de la correspondance[3].

Montrons maintenant les propriétés (1) – (4). Les propriétés (1) – (3) découlent directement du lemme 1.4. Il reste à démontrer le point (4). Soit F/K une sous-extension de L/K, et soit $H = \mathrm{Gal}(L/F)$ le sous-groupe correspondant. On a donc $F = L^H$.

Supposons que H soit distingué dans $\mathrm{Gal}(L/K)$ et soit $\tau \in \mathscr{X}(F/K)$. Alors, τ se prolonge en un plongement σ de L/K, qui est en fait un élément de $\mathrm{Gal}(L/K)$, puisque l'extension L/K est normale. Par hypothèse sur H, on a l'égalité $\sigma H \sigma^{-1} = H$. La bijectivité de la correspondance et le point (3) donnent alors

$$\tau(F) = \sigma(F) = F.$$

Cela implique que τ est un automorphisme de F/K. Ainsi, F/K est normale. Elle est aussi séparable, comme sous-extension d'une extension séparable. Par conséquent, F/K est galoisienne.

3. On peut aussi la démontrer de la même manière, grâce au lemme d'Artin.

Supposons maintenant que F/K soit galoisienne. Pour tout $\sigma \in \mathrm{Gal}(L/K)$, sa restriction $\sigma_{|F}$ à F est un plongement de F/K, donc un élément du groupe $\mathrm{Gal}(F/K)$ puisque F/K est normale. On a donc un morphisme de groupes

$$\begin{array}{rcl} \rho_F : \mathrm{Gal}(L/K) & \longrightarrow & \mathrm{Gal}(F/K) \\ \sigma & \longmapsto & \sigma_{|F}. \end{array}$$

Montrons que ρ_F est surjectif. Un élément de $\mathrm{Gal}(F/K)$ est en particulier un plongement de F/K, qui se prolonge alors en un plongement de L/K, qui est lui-même un élément de $\mathrm{Gal}(L/K)$, car L/K est normale. Cela établit la surjectivité désirée. De plus, par le lemme 1.4 (1), on a

$$\mathrm{Ker}(\rho_F) = \mathrm{Gal}(L/F).$$

En particulier, $H = \mathrm{Gal}(L/F)$ est distingué dans $\mathrm{Gal}(L/K)$, et le premier théorème d'isomorphisme appliqué à ρ_F fournit un isomorphisme

$$\mathrm{Gal}(L/K)/H \simeq \mathrm{Gal}(F/K),$$

d'où le résultat. □

1.6. Remarques. Soit L/K une extension de degré fini.

(1) On a en fait démontré que les applications $\Phi_{L/K}$ et $\Psi_{L/K}$ sont des bijections inverses l'une de l'autre si, et seulement si, L/K est une extension galoisienne.

(2) Si L/K est galoisienne, la correspondance de Galois implique que L/K admet un nombre fini de sous-extensions, puisqu'un groupe fini n'admet qu'un nombre fini de sous-groupes. C'est vrai plus généralement si l'extension L/K est monogène. En fait, une extension algébrique n'a qu'un nombre fini de sous-extensions si, et seulement si, cette extension est monogène (cf. exercice XXVII-6).

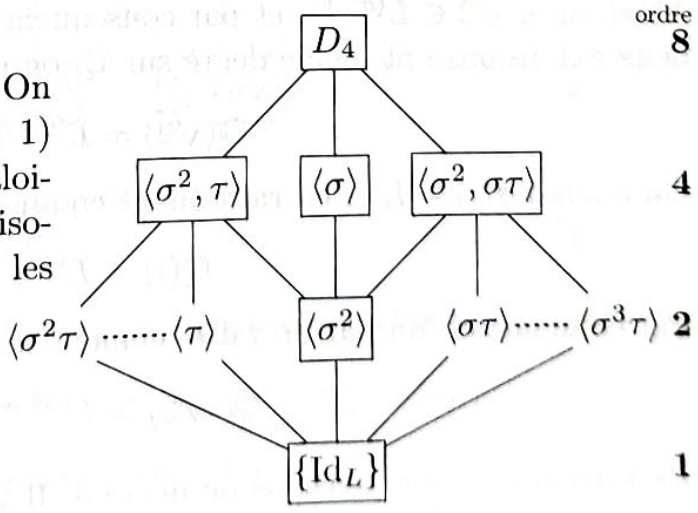
1.7. Exemple

Soit $L = \mathbb{Q}(\alpha, i)$, avec $\alpha = \sqrt[4]{2}$. On peut démontrer (cf. exercice 1) que $L/\mathbb{Q}$ est une extension galoisienne, de groupe de Galois isomorphe à D_4, engendré par les deux automorphismes

$$\sigma : \left| \begin{array}{ccc} \alpha & \rightsquigarrow & i\alpha \\ i & \rightsquigarrow & i \end{array} \right.$$

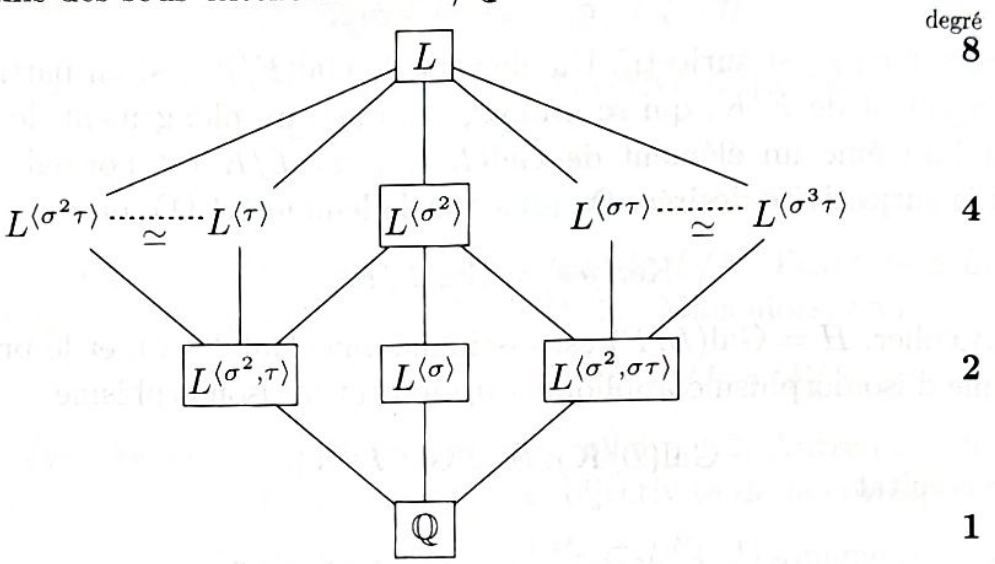
et

$$\tau : \left| \begin{array}{ccc} \alpha & \rightsquigarrow & \alpha \\ i & \rightsquigarrow & -i \end{array} \right.$$



Le treillis des sous-groupes de D_4 est donné ci-dessus, les sous-groupes distingués sont encadrés ; les pointillés relient les sous-groupes conjugués.

On a donc dix sous-groupes, les trois sous-groupes au deuxième étage sont d'indice 2, et les cinq sous-groupes au premier étage sont d'indice 4. Par la correspondance de Galois $L/\mathbb{Q}$ possède trois sous-extensions de degré 2, et cinq sous-extensions de degré 4. Cette correspondance étant décroissante, le treillis des sous-extensions de $L/\mathbb{Q}$ est de la forme



Nous allons maintenant déterminer tous ces sous-corps, et marquer ceux d'entre eux qui sont isomorphes. Remarquons que L contient $\alpha^2 = \sqrt{2}$. On a donc trois sous-extensions quadratiques « évidentes », à savoir

$$\mathbb{Q}(\sqrt{2})/\mathbb{Q}, \quad \mathbb{Q}(i)/\mathbb{Q} \quad \text{et} \quad \mathbb{Q}(i\sqrt{2})/\mathbb{Q}.$$

Pour savoir à quels sous-corps fixes elles correspondent, remarquons par exemple que l'on a

$$\sigma^2(\sqrt{2}) = \sigma^2(\alpha^2) = \sigma^2(\alpha)^2 = \big(\sigma(i\alpha)\big)^2 = (-\alpha)^2 = \sqrt{2},$$

ainsi que

$$\tau(\sqrt{2}) = \tau(\alpha^2) = \tau(\alpha)^2 = \alpha^2 = \sqrt{2}.$$

Ainsi, on a $\sqrt{2} \in L^{\langle\sigma^2,\tau\rangle}$, et par conséquent $\mathbb{Q}(\sqrt{2}) \subset L^{\langle\sigma^2,\tau\rangle}$. Comme ces deux extensions ont même degré sur $\mathbb{Q}$, on obtient

$$\mathbb{Q}(\sqrt{2}) = L^{\langle\sigma^2,\tau\rangle}.$$

On a aussi $\sigma(i) = i$, et en raisonnant comme précédemment, on obtient

$$\mathbb{Q}(i) = L^{\langle\sigma\rangle}.$$

Par élimination (ou par un raisonnement similaire), on a alors

$$\mathbb{Q}(i\sqrt{2}) = L^{\langle\sigma^2,\sigma\tau\rangle}.$$

Passons aux sous-extensions de degré 4. Il y a trois extensions auxquelles on pense immédiatement :

$$\mathbb{Q}(i\alpha)/\mathbb{Q}, \quad \mathbb{Q}(\alpha)/\mathbb{Q} \quad \text{et} \quad \mathbb{Q}(i,\sqrt{2})/\mathbb{Q}.$$

En raisonnant comme ci-dessus, on montre qu'elles correspondent respectivement aux sous-extensions

$$L^{\langle\sigma^2\tau\rangle},\quad L^{\langle\tau\rangle}\quad\text{et}\quad L^{\langle\sigma^2\rangle}.$$

Pour déterminer les deux sous-corps manquants, rappelons que

$$\mathbb{Q}(i,\sqrt{2})=\mathbb{Q}(\zeta_8).$$

Un simple calcul montre alors que l'on a

$$\sigma(\zeta_8)=-\zeta_8\quad\text{et}\quad\tau(\zeta_8)=\overline{\zeta}_8=\zeta_8^{-1}.$$

On a alors

$$\sigma\tau(\zeta_8^{-1}\alpha)=\sigma(\zeta_8\alpha)=-i\zeta_8\alpha=\zeta_8^{-1}\alpha.$$

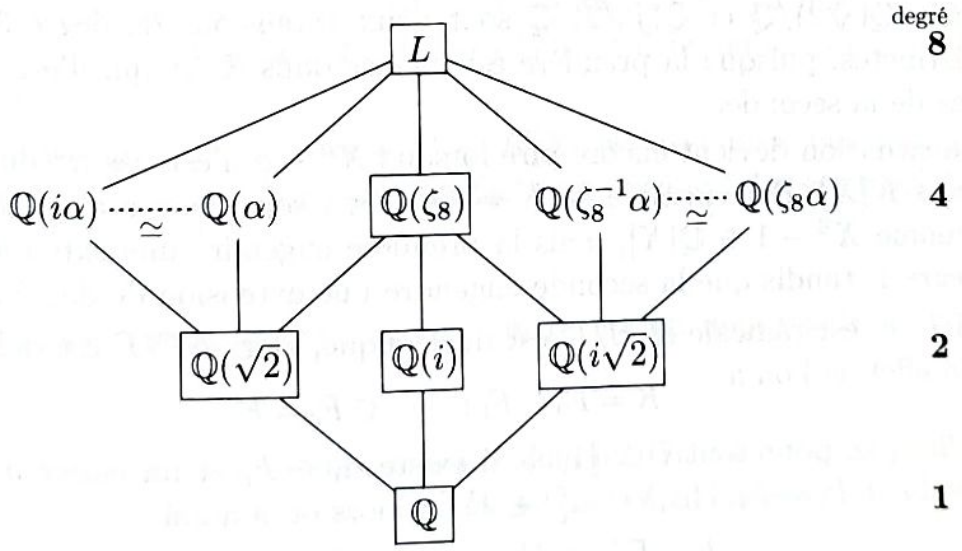
Comme précédemment, on en déduit

$$\mathbb{Q}(\zeta_8^{-1}\alpha)=L^{\langle\sigma\tau\rangle}.$$

On montre de même l'égalité $\mathbb{Q}(\zeta_8\alpha)=L^{\langle\sigma^3\tau\rangle}$.

De manière alternative, on peut remarquer que $[L:\mathbb{Q}(i,\sqrt{2})]=2$. En particulier, $(1,\alpha)$ est une $\mathbb{Q}(i,\sqrt{2})$-base de L, et l'on peut expliciter les éléments de $L^{\langle\sigma\tau\rangle}$ et de $L^{\langle\sigma^3\tau\rangle}$.

Au final, on obtient le treillis des sous-extensions de $L/\mathbb{Q}$:



Notons que les sous-groupes distingués (encadrés) de $\mathrm{Gal}(L/\mathbb{Q})$ sont

$$\mathrm{Gal}(L/\mathbb{Q}),\ \langle\sigma^2\rangle,\ \langle\sigma\rangle,\ \langle\sigma^2,\tau\rangle,\ \langle\sigma^2,\sigma\tau\rangle,\ \{\mathrm{Id}_L\},$$

et donc que seules les sous-extensions encadrées

$$L/\mathbb{Q},\ \mathbb{Q}(\zeta_8)/\mathbb{Q},\ \mathbb{Q}(i)/\mathbb{Q},\ \mathbb{Q}(\sqrt{2})/\mathbb{Q},\ \mathbb{Q}(i\sqrt{2})/\mathbb{Q},\ \mathbb{Q}/\mathbb{Q}$$

sont les sous-extensions galoisiennes de $L/\mathbb{Q}$.

2. Résolubilité des équations par radicaux

L'application que nous allons donner maintenant est l'origine historique de la théorie de Galois et de la théorie des groupes.

Avant tout, nous avons besoin de quelques définitions.

2.1. Définition. Soit K un corps de caractéristique 0. Une extension F/K est dite *radicale* s'il existe une tour d'extensions

$$K = F_0 \subset F_1 \subset \cdots \subset F_r = F$$

telle que, pour tout entier $i \in [\![1, r]\!]$, il existe $\alpha_i \in F_i$ et un entier $n_i \geqslant 1$ vérifiant $F_i = F_{i-1}(\alpha_i)$ et $\alpha_i^{n_i} \in F_{i-1}$.

2.2. Remarques

(1) Cette définition mérite quelques commentaires. Soit F/K une extension de corps. Supposons qu'il existe $\alpha \in F$ et $n \geqslant 1$ tels que $F = K(\alpha)$ et $\alpha^n \in K$. On a donc $\alpha^n = a$, pour un certain $a \in K$. On serait ainsi tenté de faire l'abus de notation $\alpha = \sqrt[n]{a}$ et $E = K(\sqrt[n]{a})$, comme dans le cas des extensions de Kummer.
La difficulté cachée est que l'on n'a pas supposé que K contient $\mu_n(\overline{K})$, si bien que chaque racine du polynôme $X^n - a$ peut engendrer une extension différente.
Par exemple, si $K = \mathbb{Q}$, $\sqrt[3]{2}$ et $j\sqrt[3]{2}$ sont deux racines de $X^3 - 2$, mais $\mathbb{Q}(\sqrt[3]{2})/\mathbb{Q}$ et $\mathbb{Q}(j\sqrt[3]{2})/\mathbb{Q}$ sont deux extensions de degré 3 bien distinctes, puisque la première est incluse dans $\mathbb{R}$, ce qui n'est pas le cas de la seconde.
La situation devient encore pire lorsque $X^n - a$ n'est pas irréductible dans $K[X]$. Par exemple, si $K = \mathbb{Q}$, 1 et i sont deux racines du polynôme $X^8 - 1 \in \mathbb{Q}[X]$, mais la première engendre une extension de degré 1, tandis que la seconde engendre une extension de degré 2.

(2) Si F/K est radicale et E/K est quelconque, alors EF/E est radicale. En effet, si l'on a

$$K = F_0 \subset F_1 \subset \cdots \subset F_r = F$$

telle que, pour tout $i \in [\![1, r]\!]$, il existe $\alpha_i \in F_i$ et un entier $n_i \geqslant 1$ vérifiant $F_i = F_{i-1}(\alpha_i)$ et $\alpha_i^{n_i} \in F_{i-1}$, alors on a aussi

$$E = EF_0 \subset EF_1 \subset \cdots \subset EF_r = EF,$$

et $EF_i = EF_{i-1}(\alpha_i)$ pour tout $i \in [\![1, r]\!]$. Ainsi, EF/E est radicale.
En particulier, le compositum d'un nombre fini d'extensions radicales est radicale. En effet, supposons que E/K et F/K soient radicales. Alors, EF/E est radicale. Mais alors, comme E/K est radicale, il est clair que EF/K est aussi radicale, puisqu'il suffit de recoller les deux tours d'extensions associées à E/K et à EF/E.

2.3. Exemples. Soit K un corps de caractéristique 0.

(1) Toute extension cyclotomique de K est radicale.

(2) Une extension de degré 2 d'un corps de caractéristique nulle est radicale d'après l'exemple XXVIII-5.3.
Par conséquent, toute extension 2-décomposable de K est radicale.

2.4. Définition. Soit K un corps de caractéristique 0. On dit que $\alpha \in \overline{K}$ est *exprimable par radicaux* s'il est contenu dans une extension E/K radicale.

2.5. Exemple. Malgré la remarque précédente, nous allons faire ici l'abus de notation usuel, à savoir que si K est un corps quelconque et $a \in K$, on désigne par $\sqrt[n]{a}$ **une** racine du polynôme $X^n - a \in K[X]$ dans $\overline{K}$.

Avec cette notation, le nombre complexe $\alpha = 2 + \sqrt[3]{7 - i\sqrt[4]{2 - \sqrt{6}}} \in \mathbb{C}$ est exprimable par radicaux.

2.6. Définition. Soit K un corps de caractéristique 0. On dit qu'un polynôme $P \in K[X]$ non constant est *résoluble par radicaux* si toutes les racines de P dans $\overline{K}$ sont exprimables par radicaux.

2.7. Remarques

(1) Puisque le compositum d'un nombre fini d'extensions radicales est radicale par la remarque 2.2 (2), un polynôme $P \in K[X]$ est résoluble par radicaux si, et seulement si, $\mathrm{Dec}_K(P)$ est contenu dans une extension radicale E/K.

(2) Clairement, un polynôme $P \in K[X]$ est résoluble par radicaux si, et seulement si, chacun de ses facteurs irréductibles est résoluble par radicaux.

Nous allons maintenant introduire l'autre notion importante de ce paragraphe.

2.8. Définition. Soit K un corps quelconque. On dit qu'une extension de corps L/K de degré fini est *résoluble* si elle est galoisienne, et $\mathrm{Gal}(L/K)$ est un groupe résoluble.

2.9. Exemples

(1) Une extension abélienne est résoluble, puisqu'un groupe abélien est résoluble.

(2) Si L/K est galoisienne, de degré p^n, où p est premier, alors L/K est résoluble, car un p-groupe fini est résoluble par le théorème IX-3.23.

2.10. Lemme. *Soit K un corps. Soit L/K une extension galoisienne, et soit F/K une sous-extension de L/K. Alors :*

(1) *si L/K est résoluble, alors L/F est résoluble ;*

(2) *si L/K est résoluble et F/K est galoisienne, alors F/K est résoluble ;*

(3) *si F/K est galoisienne, alors L/K est résoluble si, et seulement si, les extensions F/K et L/F sont résolubles ;*

(4) *si L/K est une extension résoluble, et si E/K est une extension quelconque, alors EL/E est résoluble ;*

(5) *le compositum d'un nombre fini d'extensions résolubles de K est une extension résoluble de K.*

Démonstration. Soit L/K une extension galoisienne, et soit $G = \mathrm{Gal}(L/K)$. Alors, $H = \mathrm{Gal}(L/F)$ est un sous-groupe de G. Si de plus, F/K est galoisienne, alors H est distingué dans G par la correspondance de Galois, et l'on a dans ce cas $\mathrm{Gal}(F/K) \simeq G/H$. Les propriétés (1) – (3) découlent alors directement de la proposition IX-3.21.

Montrons (4). Si L/K est résoluble et E/K est une extension quelconque, la proposition XXVIII-2.5 montre que LE/E est galoisienne, et que $\mathrm{Gal}(LE/E)$ est isomorphe à un sous-groupe de $\mathrm{Gal}(L/K)$. Comme $\mathrm{Gal}(L/K)$ est résoluble, $\mathrm{Gal}(LE/E)$ l'est également.

Il reste à vérifier (5). Il suffit de montrer que le compositum de deux extensions résolubles est encore résoluble. Soient L/K et L'/K deux telles extensions. D'après le point (4), l'extension LL'/L' est résoluble. Comme L'/K est résoluble, le point (3) montre que LL'/K est résoluble. Cela achève la démonstration. □

Le lemme suivant relie les notions d'élément exprimable par radicaux et d'extension résoluble.

2.11. Lemme. *Soit K un corps de caractéristique 0. Alors, tout élément d'une extension résoluble est exprimable par radicaux.*

Démonstration. Soit L/K une extension résoluble de degré $n \geqslant 1$.

Soit $E = K(\zeta_n)$. L'extension E/K étant résoluble, EL/E est résoluble par le lemme précédent. D'après la proposition XXVIII-2.5, $\mathrm{Gal}(EL/E)$ est isomorphe à un sous-groupe de $\mathrm{Gal}(L/K)$, qui est d'ordre n puisque L/K est galoisienne. On a donc

$$|\mathrm{Gal}(EL/E)| \mid n.$$

Puisque E/K est radicale et que EL contient les éléments de L, il suffit de montrer que EL/E est radicale.

On peut donc supposer sans perte de généralité que $\mu_n(\overline{K}) \subset K$ et que l'extension L/K est résoluble, de degré divisant n.

Posons $G = \mathrm{Gal}(L/K)$. Puisque G est résoluble, il existe une suite de composition

$$G_r = \{\mathrm{Id}_L\} \lhd G_{r-1} \lhd \cdots \lhd G_1 \lhd G_0 = G,$$

avec G_i/G_{i+1} abélien pour tout $i \in [\![0, r-1]\!]$.

Posons $F_i = L^{G_i}$ pour tout $i \in [\![0, r]\!]$. Ainsi, on a

$$K = F_0 \subset F_1 \subset \cdots \subset F_r = L.$$

Soit $i \in [\![0, r-1]\!]$. Le lemme d'Artin montre que

$$\mathrm{Gal}(L/F_{i+1}) = G_{i+1} \lhd G_i = \mathrm{Gal}(L/F_i).$$

D'après la correspondance de Galois, l'extension F_{i+1}/F_i est galoisienne, et $\mathrm{Gal}(F_{i+1}/F_i) \simeq G_i/G_{i+1}$ est un groupe abélien. Remarquons enfin que l'entier $[F_{i+1} : F_i]$ divise $[L : K]$, et donc divise n.

Il suffit donc de démontrer que chaque extension F_{i+1}/F_i est radicale. Autrement dit, on peut supposer que $\mu_n(\overline{K}) \subset K$, et que L/K est une extension abélienne de degré divisant n, ce que l'on fera désormais.

Le théorème XII-5.3 montre alors qu'il existe des groupes cycliques $H_1, \ldots, H_s$ et un isomorphisme de groupes

$$\varphi : G \xrightarrow{\sim} H_1 \times \cdots \times H_s.$$

Pour tout $j \in [\![0, s]\!]$, posons

$$G_j = \varphi^{-1}(H_1 \times \cdots \times H_{s-j} \times \{1\} \times \cdots \times \{1\}).$$

Alors, pour tout $j \in [\![0, s-1]\!]$, G_{j+1} est distingué dans G_j, et l'on a

$$G_j/G_{j+1} \simeq H_{s-j}.$$

Le même argument que précédemment nous permet alors de nous réduire au cas d'une extension cyclique dont le degré divise n (avec $\mu_n(\overline{K}) \subset K$). La théorie de Kummer implique qu'une telle extension est radicale. Cela achève la démonstration. □

Nous allons maintenant nous intéresser à la démonstration de la réciproque de ce résultat, à savoir qu'un élément exprimable par radicaux est contenu dans une extension résoluble. On commence par un lemme technique.

2.12. Lemme. *Soit K un corps de caractéristique 0, et soit E/K une extension galoisienne. Soit F/E une extension telle qu'il existe $\alpha \in F$ et un entier $n \geqslant 1$ vérifiant $F = E(\alpha)$ et $\alpha^n \in E$. Alors, F est contenue dans une extension galoisienne L/K telle que L/E soit résoluble. Si de plus $n = 2$, on peut choisir L/K telle que $[L : E]$ soit une puissance de 2.*

Démonstration. Soit $a \in E$ tel que $\alpha^n = a$, et posons

$$P = \prod_{\tau \in \mathrm{Gal}(E/K)} \big(X^n - \tau(a)\big) \in E[X].$$

Pour tout $\sigma \in \mathrm{Gal}(E/K)$, on a

$$\sigma(P) = \prod_{\tau \in \mathrm{Gal}(E/K)} \big(X^n - \sigma\tau(a)\big) = \prod_{\tau \in \mathrm{Gal}(E/K)} \big(X^n - \tau(a)\big).$$

Ainsi, les coefficients de P sont fixés par tous les éléments de $\mathrm{Gal}(E/K)$, donc ce sont des éléments de K. Autrement dit, $P \in K[X]$.

Soient $\alpha = \alpha_1, \ldots, \alpha_m \in \overline{K}$ les racines de P. Posons

$$M = \mathrm{Dec}_K(P) = K(\alpha_1, \ldots, \alpha_m).$$

Alors, M/K est normale par construction, et séparable puisque K est de caractéristique nulle. Ainsi, M/K est galoisienne. Posons $L = K(\zeta_n)EM$. Comme $E/K, M/K$ et $K(\zeta_n)/K$ sont galoisiennes, L/K est aussi galoisienne par le lemme XXVIII-2.9. Par construction, $F = E(\alpha) \subset L$.

Montrons que L/E est résoluble. Remarquons que $L/E(\zeta_n)$ est le compositum des extensions $E(\zeta_n, \alpha_i)/E(\zeta_n)$, $i \in [\![1, m]\!]$. Or, par définition de P et des α_i, pour tout $i \in [\![1, m]\!]$, il existe $a_i \in E$ tel que $\alpha_i^n = a_i$. Par conséquent, l'extension $E(\zeta_n, \alpha_i)/E(\zeta_n)$ est une extension de Kummer, donc cyclique, et en particulier résoluble. Le dernier point du lemme 2.10 assure alors que $L/E(\zeta_n)$ est résoluble. Comme $E(\zeta_n)/E$ est une extension abélienne, donc résoluble, le point (3) de ce même lemme entraîne que L/E est résoluble.

Supposons maintenant que $n = 2$, et vérifions que $[L : E]$ est une puissance de 2. Remarquons que dans ce cas, on a $L = E(\alpha_1, \ldots, \alpha_n)$. Remarquons aussi que chaque α_i est une racine d'un polynôme de $E[X]$ de degré 2, et est donc de degré au plus 2 sur E. Or, on a le fait suivant.

Fait. Soit E'/E une extension de degré fini, et soit $\alpha \in \overline{E}$ de degré au plus 2 sur E. Alors, $[E'(\alpha) : E] = [E' : E]$ ou $2[E' : E]$.

En effet, on a $[E'(\alpha) : E'] \leqslant [E(\alpha) : E] \leqslant 2$ par le lemme XXV-3.20. On a donc $[E'(\alpha) : E'] = 1$ ou 2, d'où le résultat par la formule de multiplicativité des degrés.

En utilisant ce fait, on démontre par récurrence que $[E(\alpha_1, \ldots, \alpha_n) : E]$ est une puissance de 2. Cela achève la démonstration. □

On en déduit le résultat suivant.

2.13. Lemme. *Soit K un corps de caractéristique 0. Alors, toute extension radicale F/K est contenue dans une extension résoluble L/K.*

Si de plus F/K est 2-décomposable, on peut choisir L/K telle que $[L : K]$ soit une puissance de 2.

Démonstration. Rappelons qu'une extension 2-décomposable est radicale, d'après l'exemple 2.3 (2). Nous allons démontrer la propriété suivante pour tout $r \geqslant 0$ par récurrence :

(H_r) pour toute extension F/K radicale (resp. 2-décomposable) telle que

$$K = F_0 \subset F_1 \subset \cdots \subset F_{r-1} \subset F_r = F,$$

avec $F_i = F_{i-1}(\alpha_i)$, où $\alpha_i \in F_i$ vérifie $\alpha_i^{n_i} \in F_{i-1}$ pour un certain $n_i \geqslant 1$ (resp. $n_i \leqslant 2$) pour tout $i \in [\![1, r]\!]$, l'extension F est contenue dans une extension résoluble de K (resp. une extension galoisienne dont le degré est une puissance de 2).

La propriété (H_0) étant trivialement vraie, supposons que (H_r) soit vraie pour un certain entier $r \geqslant 0$, et considérons une extension radicale (resp.2-décomposable) F/K de la forme

$$K = F_0 \subset F_1 \subset \cdots \subset F_{r-1} \subset F_r \subset F_{r+1} = F$$

vérifiant les propriétés adéquates.

Par hypothèse de récurrence, F_r/K est une sous-extension d'une extension résoluble E/K. De plus, on peut supposer que $[E : K]$ est une puissance de 2 dans le cas 2-décomposable.

Puisque E contient F_r, on a $E = EF_r$. On a alors $EF_{r+1} = E(\alpha_{r+1})$. D'après le lemme précédent, EF_{r+1} est contenue dans une extension galoisienne L/K telle que L/E soit résoluble. Dans le cas 2-décomposable, on peut supposer de plus que $[L : E]$ est une puissance de 2. Comme les extensions E/K et L/E sont résolubles, l'extension L/K est résoluble par le lemme 2.10. Dans le cas 2-décomposable, $[E : K]$ et $[L : E]$ sont des puissances de 2, et il en est donc de même de $[L : K]$. Cela achève la récurrence, ainsi que la démonstration. □

On peut maintenant énoncer le théorème suivant.

2.14. Théorème. *Soit K un corps de caractéristique 0, et soit $\alpha \in \overline{K}$. Les propriétés suivantes sont équivalentes :*

(1) *α est exprimable par radicaux ;*

(2) *α est contenu dans une extension résoluble de K ;*

(3) *le groupe $\mathrm{Gal}_K(\mu_{\alpha,K})$ est résoluble.*

Démonstration

(1) $\Longrightarrow$ (2). Si α est exprimable par radicaux, il est contenu dans une extension F/K radicale. Le lemme 2.13 montre alors que F/K est une

sous-extension d'une extension L/K résoluble. Comme $\alpha \in L$, on obtient le résultat voulu.

(2) $\Longrightarrow$ (3). Soit L/K une extensions résoluble contenant α. Puisque $\alpha \in F$, l'extension normale L/K contient une racine de $\mu_{\alpha,K}$, et donc les contient toutes. Ainsi, on a

$$\mathrm{Dec}_K(\mu_{\alpha,K}) \subset L.$$

Comme L/K est résoluble et $\mathrm{Dec}_K(\mu_{\alpha,K})/K$ est galoisienne, cette dernière extension est aussi résoluble par le lemme 2.10 (2).

(3) $\Longrightarrow$ (1). Par hypothèse, $\mathrm{Dec}_K(\mu_{\alpha,K})/K$ est résoluble et contient α. D'après le lemme 2.11, α est exprimable par radicaux.

Cela achève la démonstration. □

2.15. Corollaire. *Soit K un corps de caractéristique 0, et soit $P \in K[X]$ un polynôme non constant. Les propriétés suivantes sont équivalentes :*

(1) *le polynôme P est résoluble par radicaux ;*

(2) $\mathrm{Dec}_K(P)$ *est contenu dans une extension résoluble de K ;*

(3) *le groupe* $\mathrm{Gal}_K(P)$ *est résoluble.*

Démonstration

(1) $\Longrightarrow$ (2). Soient $\alpha_1, \ldots, \alpha_n \in \overline{K}$ les racines de P. D'après le théorème précédent, chaque α_i est contenu dans une extension résoluble L_i/K. Posons $L = L_1 \ldots L_n$. Alors, L/K est résoluble par le lemme 2.10 (5). Comme L/K contient toutes les racines de P, elle contient $\mathrm{Dec}_K(P)$.

(2) $\Longrightarrow$ (3). Soit L/K une extension résoluble contenant $\mathrm{Dec}_K(P)$. Comme l'extension $\mathrm{Dec}_K(P)/K$ est galoisienne, cette dernière extension est aussi résoluble par le lemme 2.10 (2).

(3) $\Longrightarrow$ (1). Comme $\mathrm{Dec}_K(P)/K$ est résoluble et contient toutes les racines de P, il suffit d'appliquer le théorème précédent pour conclure que P est résoluble par radicaux. □

2.16. Exemple. Un polynôme de degré $\geqslant 5$ à coefficients dans $\mathbb{Q}$ n'est pas en général résoluble par radicaux. Ce résultat, dû au mathématicien norvégien Abel[4], est antérieur aux résultats précédents, qui eux ont été démontrés par Galois.

Pour le voir, il suffit d'exhiber un exemple de polynôme irréductible à coefficients rationnels de degré n dont le groupe de Galois n'est pas résoluble. Or, pour tout $n \geqslant 5$, on peut construire un polynôme de $\mathbb{Q}[X]$ dont le groupe de Galois est isomorphe à $\mathfrak{S}_n$ (cf. exercice XXVIII-19).

4. À ne pas confondre avec l'Abel de Cadix, qui lui a des yeux de velours.

Par exemple, le polynôme $X^5 - X - 1 \in \mathbb{Q}[X]$ n'est pas résoluble par radicaux, puisque son groupe de Galois est isomorphe à $\mathfrak{S}_5$, qui n'est pas résoluble (cf. exemple XXVIII-3.11).

Il est en revanche bien connu qu'un polynôme de degré $\leqslant 4$ est résoluble par radicaux. C'est évident pour les degrés 1 et 2. Pour le degré 3, la méthode de Cardan permet de se ramener à la résolution d'une équation de degré 2. Pour le degré 4, la méthode de Ferrari permet quant à elle de se ramener au cas d'une équation de degré 3.

Cela peut se retrouver aisément grâce aux résultats de cette section. En effet, soit K un corps de caractéristique 0, et soit $P \in K[X]$ un polynôme de degré $n \leqslant 4$. Comme $\mathrm{Gal}_K(P)$ est isomorphe à un sous-groupe de $\mathfrak{S}_n$, qui est résoluble, $\mathrm{Gal}_K(P)$ est résoluble. On conclut en utilisant le corollaire précédent.

3. Constructions à la règle et au compas : le retour

Dans cette courte section, nous donnons une condition nécessaire et suffisante sur la constructibilité d'un nombre complexe algébrique à la règle et au compas en termes de groupe de Galois.

3.1. Théorème. *Soit $\mathscr{A}_0$ un ensemble de points de $\mathbb{R}^2$ contenant $O = (0,0)$ et $I = (1,0)$ et soit $\alpha \in \mathbb{C}$. Les propriétés suivantes sont équivalentes :*

(1) *α est $\mathscr{A}_0$-constructible ;*

(2) *α est contenu dans une extension galoisienne $L/\mathbb{Q}(\mathscr{A}_0)$ dont le degré est une puissance de 2 ;*

(3) *α est algébrique sur $\mathbb{Q}(\mathscr{A}_0)$ et $\mathrm{Gal}_{\mathbb{Q}(\mathscr{A}_0)}(\mu_{\alpha,\mathbb{Q}(\mathscr{A}_0)})$ est un 2-groupe.*

Démonstration. Pour alléger les notations, posons $K = \mathbb{Q}(\mathscr{A}_0)$.

(1) $\Longrightarrow$ (2). Si α est un nombre complexe $\mathscr{A}_0$-constructible, il est contenu dans une extension F/K 2-décomposable. Le lemme 2.13 montre alors que F/K est une sous-extension d'une extension galoisienne L/K dont le degré est une puissance de 2. Comme $\alpha \in F$, on obtient le résultat voulu.

(2) $\Longrightarrow$ (3). Puisque α est contenu dans une extension de K de degré fini, il est algébrique sur K. De plus, puisque $\alpha \in L$, l'extension normale L/K contient une racine de $\mu_{\alpha,K}$, et donc les contient toutes.

Ainsi, on a

$$\mathrm{Dec}_K(\mu_{\alpha,K}) \subset L.$$

Comme $[\mathrm{Dec}_K(\mu_{\alpha,K}) : K]$ divise $[L : K]$, $[\mathrm{Dec}_K(\mu_{\alpha,K}) : K]$ est aussi une puissance de 2. Comme l'extension $\mathrm{Dec}_K(\mu_{\alpha,K})/K$ est galoisienne, on en déduit que $\mathrm{Gal}_K(\mu_{\alpha,K})$ est un 2-groupe.

(3) $\Longrightarrow$ (1). Soit $G = \mathrm{Gal}_K(\mu_{\alpha,K})$. Comme G est un 2-groupe, d'après le théorème IX-3.23, il existe une suite de composition

$$G_r = \{\mathrm{Id}\} \lhd G_{r-1} \lhd \cdots \lhd G_1 \lhd G_0 = G,$$

où $[G_i : G_{i+1}] = 2$ pour tout $i \in [\![0, r-1]\!]$. En posant $F_i = \mathrm{Dec}_K(\mu_{\alpha,K})^{G_i}$, on obtient une tour d'extensions

$$K \subset F_0 \subset \cdots \subset F_{r-1} \subset F_r = \mathrm{Dec}_K(\mu_{\alpha,K}).$$

De plus, par le lemme d'Artin, on a $[F_{i+1} : F_i] = [G_i : G_{i+1}] = 2$ pour tout entier $i \in [\![0, r-1]\!]$. Ainsi, l'extension $\mathrm{Dec}_K(\mu_{\alpha,K})/K$ est 2-décomposable et contient α. Cela achève la démonstration. □

3.2. Remarque. Ce théorème permet de retrouver de façon plus directe que ζ_n est constructible si, et seulement si, $\varphi(n)$ est une puissance de 2. En effet, on a $\mu_{\zeta_n,\mathbb{Q}} = \Phi_n$, et $\mathrm{Dec}_{\mathbb{Q}}(\Phi_n) = \mathbb{Q}(\zeta_n)$, qui est de degré $\varphi(n)$. On a donc

$$|\mathrm{Gal}(\mu_{\zeta_n,\mathbb{Q}})| = \varphi(n),$$

et l'on applique alors le théorème 3.1.

3.3. Exemple. Soit $P = X^4 - X - 1 \in \mathbb{Q}[X]$, et soit X_P l'ensemble des racines complexes de P.

On vérifie que P est irréductible modulo 2. En effet, $\overline{P}$ n'a pas de racines dans $\mathbb{F}_2$, et n'est pas divisible par l'unique polynôme unitaire de degré 2 irréductible de $\mathbb{F}_2[X]$, à savoir $X^2 + X + \overline{1}$. De manière alternative, on peut utiliser la proposition XXI-6.14, et remarquer que P n'a pas de racines dans $\mathbb{F}_2$ et dans $\mathbb{F}_4$, puisque si $x \in \mathbb{F}_4$, on a $x^4 - x - 1 = -1 \neq 0$.

Soit $\alpha \in X_P$. On a donc $\mu_{\alpha,\mathbb{Q}} = P$. Remarquons maintenant que

$$\overline{P} = (X + \overline{4})(X^3 + \overline{3}X^2 + \overline{2}X + \overline{5}) \in \mathbb{F}_7[X].$$

Le deuxième facteur est irréductible, car il n'a pas de racines dans $\mathbb{F}_7$. D'après le théorème XXVIII-3.10, le groupe $\mathrm{Gal}_{\mathbb{Q}}(P)$ (identifié à un sous-groupe de $\mathfrak{S}(X_P)$) contient un 3-cycle. En particulier, $3 \mid |\mathrm{Gal}_{\mathbb{Q}}(P)|$. On en déduit que α n'est pas constructible à la règle et au compas.

On peut même préciser ici le groupe de Galois. L'irréductibilité de P modulo 2 montre que $\mathrm{Gal}_{\mathbb{Q}}(P)$ contient un 4-cycle. Ainsi, $4 \mid |\mathrm{Gal}_{\mathbb{Q}}(P)|$, et par conséquent $12 \mid |\mathrm{Gal}_{\mathbb{Q}}(P)|$. Comme $\mathrm{Gal}_{\mathbb{Q}}(P)$ s'identifie à un sous-groupe de $\mathfrak{S}(X_P)$, qui est d'ordre 24, $\mathrm{Gal}_{\mathbb{Q}}(P)$ est donc d'ordre 12 ou 24. Il est donc isomorphe à $\mathfrak{A}(X_P)$ ou $\mathfrak{S}(X_P)$ par la proposition VIII-4.8. Or, on

vérifie que P a exactement deux racines réelles, et par conséquent deux racines complexes conjuguées. La conjugaison complexe induit alors une transposition de $\mathfrak{S}(X_P)$. Cela exclut la première possibilité. Finalement, on a

$$\mathrm{Gal}_{\mathbb{Q}}(P) \simeq \mathfrak{S}(X_P) \simeq \mathfrak{S}_4.$$

3.4. Corollaire. *Soit $\mathscr{A}_0$ un ensemble de points de $\mathbb{R}^2$ contenant les deux points $O = (0,0)$ et $I = (1,0)$, et soit $P \in \mathbb{Q}(\mathscr{A}_0)[X]$. Les propriétés suivantes sont équivalentes :*

(1) *toute racine de P est $\mathscr{A}_0$-constructible ;*

(2) $\mathrm{Dec}_{\mathbb{Q}(\mathscr{A}_0)}(P)$ *est contenu dans une extension de $\mathbb{Q}(\mathscr{A}_0)$ dont le degré est une puissance de 2 ;*

(3) *le groupe $\mathrm{Gal}_{\mathbb{Q}(\mathscr{A}_0)}(P)$ est un 2-groupe.*

Démonstration. Posons $K = \mathbb{Q}(\mathscr{A}_0)$.

(1) $\Longrightarrow$ (2). Soient $\alpha_1, \ldots, \alpha_n \in \mathbb{C}$ les racines complexes de P. Par hypothèse, chaque α_i est $\mathscr{A}_0$-constructible, donc contenu dans une extension galoisienne L_i/K dont le degré est une puissance de 2. Posons

$$L = L_1 \cdots L_n.$$

Alors, l'extension L/K est galoisienne. De plus, le lemme XXVIII-2.9 et une récurrence facile montrent que le groupe $\mathrm{Gal}(L/K)$ s'identifie à un sous-groupe de $\mathrm{Gal}(L_1/K) \times \cdots \times \mathrm{Gal}(L_n/K)$, qui est un 2-groupe. En particulier, $\mathrm{Gal}(L/K)$ est un 2-groupe et $[L : K]$ est une puissance de 2. Comme L/K contient toutes les racines de P, elle contient aussi $\mathrm{Dec}_K(P)$.

(2) $\Longrightarrow$ (3). Soit L/K une extension galoisienne dont le degré est une puissance de 2, et contenant $\mathrm{Dec}_K(P)$. Comme $\mathrm{Dec}_K(P)/K$ est galoisienne, son groupe de Galois est isomorphe à un quotient de $\mathrm{Gal}(L/K)$ par la correspondance de Galois, donc est aussi un 2-groupe.

(3) $\Longrightarrow$ (1). Comme $\mathrm{Dec}_K(P)/K$ est galoisienne de degré une puissance de 2 et contient toutes les racines de P, il suffit d'appliquer le théorème précédent pour conclure. □

4. Exercices

E1. Soit p un nombre premier, soit $P = X^4 - p \in \mathbb{Q}[X]$ et soit $L = \mathbb{Q}(\alpha, i)$, où $\alpha = \sqrt[4]{p}$.

a. Vérifier que P est irréductible et que $L/\mathbb{Q}$ est galoisienne.

b. Justifier que $[L : \mathbb{Q}(\alpha)] = 2$. En déduire que $[L : \mathbb{Q}] = 8$.

c. Montrer que $\mathbb{Q}(\alpha)/\mathbb{Q}$ n'est pas galoisienne.

d. Sans absolument aucun calcul, identifier $\mathrm{Gal}(L/\mathbb{Q})$.

Indication. Quels sont les groupes G d'ordre 8 qui possèdent un sous-groupe non distingué dans G ?

e. Calculer les éléments de $\mathrm{Gal}(L/\mathbb{Q})$ et retrouver le résultat de la question précédente.

E2. Soit $P = X^8 - 3 \in \mathbb{Q}[X]$, et soit $L = \mathrm{Dec}_{\mathbb{Q}}(P)$.

a. On pose $\theta = \sqrt[8]{3}$. Justifier que $L = \mathbb{Q}(\theta, \zeta_8)$.

On pose $K = \mathbb{Q}(\theta)$ et $k = \mathbb{Q}(\zeta_8) = \mathbb{Q}(i, \sqrt{2})$.

b. Montrer que $K \cap k = \mathbb{Q}$ ou $\mathbb{Q}(\sqrt{2})$.

c. En utilisant la théorie de Kummer, montrer que $X^8 - 3$ est irréductible sur k. En déduire $[L : \mathbb{Q}]$ et $K \cap k$.

d. Montrer que l'extension $K/\mathbb{Q}$ n'est pas galoisienne. Étudier les deux extensions $\mathbb{Q}(i, \theta^2)/\mathbb{Q}$ et $\mathbb{Q}(i, \theta)/\mathbb{Q}$.

Montrer que $\mathrm{Gal}(L/\mathbb{Q})$ est isomorphe à $\mathbb{Z}/8\mathbb{Z} \times_\rho (\mathbb{Z}/2\mathbb{Z})^2$. On précisera le morphisme $\rho : (\mathbb{Z}/2\mathbb{Z})^2 \longrightarrow \mathrm{Aut}(\mathbb{Z}/8\mathbb{Z})$.

E3. Le but de cet exercice est de démontrer que tout groupe abélien fini G peut se réaliser comme le groupe de Galois d'une extension galoisienne de $\mathbb{Q}$. On suppose G non trivial, et l'on écrit

$$G \simeq \mathbb{Z}/d_1\mathbb{Z} \times \cdots \times \mathbb{Z}/d_r\mathbb{Z}$$

avec $d_1 \mid \cdots \mid d_r$ $(d_i \geqslant 2)$.

a. Montrer qu'il existe des nombres premiers $p_1, \ldots, p_r$ deux à deux distincts tels que $p_i \equiv 1\ [d_i]$ pour tout $i \in [\![1, r]\!]$.

b. Montrer que pour tout $i \in [\![1, r]\!]$, il existe une sous-extension galoisienne $L_i/\mathbb{Q}$ de $\mathbb{Q}(\zeta_{p_i})/\mathbb{Q}$ telle que $\mathrm{Gal}(L_i/\mathbb{Q}) \simeq \mathbb{Z}/d_i\mathbb{Z}$.

c. Montrer que $(L_1 \cdots L_{i-1}) \cap L_i = \mathbb{Q}$ pour tout $i \in [\![1, r]\!]$. En déduire que $L_1 \cdots L_{i-1}/\mathbb{Q}$ et $L_i/\mathbb{Q}$ sont linéairement disjointes.

Indication. Que vaut $\big(\mathbb{Q}(\zeta_{p_1}) \cdots \mathbb{Q}(\zeta_{p_{i-1}})\big) \cap \mathbb{Q}(\zeta_{p_i})$?

d. Soit $L = L_1 \cdots L_r$. Montrer que $\mathrm{Gal}(L/\mathbb{Q}) \simeq G$.

E4. Soit K un corps de caractéristique différente de 2, soit $n \geqslant 1$ un entier et soient $a_1, \ldots, a_n \in K^\times$.

Pour tout sous-ensemble **non vide** I de $[\![1, n]\!]$, on pose

$$a_I = \prod_{i \in I} a_i.$$

On dit que $a_1, \ldots, a_n$ sont *quadratiquement indépendants sur* K si pour tout I non vide, $a_I \in K^\times \setminus K^{\times 2}$.

On pose

$$L_n = K\big(\sqrt{a_1}, \ldots, \sqrt{a_n}\big).$$

a. Montrer que L_n/K est galoisienne, et que $[L_n : K] \leqslant 2^n$.

Le but de ce qui suit est entre autres de démontrer que, pour tous éléments $a_1, \ldots, a_n \in K^\times$ quadratiquement indépendants sur K, L_n/K est une extension galoisienne de groupe de Galois isomorphe à $(\mathbb{Z}/2\mathbb{Z})^n$. On procède par récurrence sur $n \geqslant 1$.

b. Démontrer le résultat pour $n = 1$ et $n = 2$.

On suppose que le résultat est vrai au rang n.

c. Si $a_1, \ldots, a_n \in K^\times$ sont quadratiquement indépendants sur K, montrer que les sous-extensions de degré 2 de L_n/K sont exactement les extensions $K(\sqrt{a_I})/K$, lorsque I parcourt l'ensemble des parties non vides de $[\![1, n]\!]$.

Indication. Vérifier que $(\mathbb{Z}/2\mathbb{Z})^n$ possède exactement $2^n - 1$ sous-groupes d'indice 2.

d. On suppose que $a_1, \ldots, a_{n+1} \in K^\times$ sont quadratiquement indépendants sur K.

(i) Montrer que $[L_{n+1} : L_n] \leqslant 2$.

(ii) On suppose que $L_{n+1} = L_n$. Montrer qu'il existe une partie I non vide de $[\![1, n]\!]$ telle que

$$K(\sqrt{a_{n+1}}) = K(\sqrt{a_I}),$$

et en déduire une contradiction.

(iii) En déduire que L_n/K et $K(\sqrt{a_{n+1}})/K$ sont linéairement disjointes, et conclure.

e. Montrer que les éléments de $\mathrm{Gal}(L_n/K)$ sont les automorphismes

$$\sigma_\varepsilon : \sqrt{a_i} \rightsquigarrow \varepsilon_i \sqrt{a_i},$$

où $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n) \in \{\pm 1\}^n$.

f. Montrer que $\sqrt{a_1} + \cdots + \sqrt{a_n}$ est un élément primitif de L_n/K.

g. Soient $p_1, \ldots, p_n$ des nombres premiers deux à deux distincts. Quel est le groupe de Galois de $\mathbb{Q}(\sqrt{p_1}, \ldots, \sqrt{p_n})/\mathbb{Q}$?

E5. Soit $n \geqslant 2$.

a. Soit $H = \{\sigma \in \mathfrak{S}_n \mid \sigma(n) = n\}$. Montrer que H est un sous-groupe maximal de $\mathfrak{S}_n$ d'indice n.

b. Montrer que pour tout $n \geqslant 2$, il existe une extension $K/\mathbb{Q}$ de degré n sans sous-corps intermédiaires non triviaux (i.e. différents de K et $\mathbb{Q}$).

Indication. On pourra utiliser une extension galoisienne $L/\mathbb{Q}$ de groupe de Galois isomorphe à $\mathfrak{S}_n$.

E6. Soit $P = X^6 - 3X^2 - 1 \in \mathbb{Q}[X]$ et soit $Q = X^3 - 3X - 1 \in \mathbb{Q}[X]$. On note $L = \mathrm{Dec}_{\mathbb{Q}}(P)$ et $G = \mathrm{Gal}_{\mathbb{Q}}(P)$.

a. Montrer que P et Q sont irréductibles. Vérifier aussi que Q a une racine positive α et deux racines négatives β et γ.

b. Montrer que $\beta = -\dfrac{1}{1+\alpha}$ et $\gamma = -1 - \dfrac{1}{\alpha}$.

c. Soit $K = \mathbb{Q}(\alpha)$. Montrer que $\mathbb{Q}(\alpha)/\mathbb{Q}$ est galoisienne de degré 3, et que $\mathrm{Gal}(K/\mathbb{Q})$ est engendré par

$$\sigma : \alpha \rightsquigarrow \beta.$$

Vérifier que $\psi(\beta) = \gamma$.

d. Montrer que $[K(\sqrt{\alpha}) : K] = 2$ et $[K(\sqrt{\beta}) : K] = 2$.
Indication. Si α, $\beta \in K^{\times 2}$, que dire de leurs images par σ ?

e. Montrer que $K(\sqrt{\alpha})/K$ et $K(\sqrt{\beta})/K$ sont linéairement disjointes, et que $\mathrm{Gal}\big(K(\sqrt{\alpha}, \sqrt{\beta})/K\big) \simeq (\mathbb{Z}/2\mathbb{Z})^2$.

f. Vérifier que le corps des racines de P est $L = \mathbb{Q}(\sqrt{\alpha}, \sqrt{\beta})$.

g. En utilisant les questions précédentes, montrer que G est d'ordre 12, et qu'un 2-Sylow est isomorphe à $(\mathbb{Z}/2\mathbb{Z})^2$ et distingué dans G.

h. Calculer explicitement les éléments de G, ainsi que leurs ordres.

i. En utilisant l'exercice XI-10, identifier le groupe G.

E7. Soit K un corps.

a. Montrer que le compositum de deux extensions 2-décomposables de K est encore 2-décomposable.

b. Soit M/K une extension séparable 2-décomposable. Montrer que toute sous-extension F/K de M/K est 2-décomposable.

Indication. Considérer une extension galoisienne L/K contenant M, dont le groupe de Galois est un 2-groupe, et utiliser l'exercice IX-16.

c. Soit $\mathscr{A}_0$ un ensemble fini de points de $\mathbb{R}^2$ contenant $(0,0)$ et $(1,0)$. Montrer $\alpha \in \mathbb{C}$ est $\mathscr{A}_0$-constructible si, et seulement si, $\mathbb{Q}(\mathscr{A}_0)(\alpha)/\mathbb{Q}(\mathscr{A}_0)$ est 2-décomposable.

E8. Soit une extension F/K contenue dans une extension galoisienne dont le degré est une puissance de 2. En utilisant l'exercice XII-5, montrer que F/K est séparable et 2-décomposable.

E9. On dit qu'une extension L/K est $\{2,3\}$-décomposable s'il existe une tour d'extensions

$$E_0 = K \subset E_1 \subset \cdots \subset E_r = L$$

vérifiant $[E_i : E_{i-1}] \in \{1,2,3\}$ pour tout $i \in [\![1, r]\!]$.

On admet que tout groupe fini d'ordre $p^m q^m$, où p, q sont des nombres premiers, est résoluble (cf. théorème XXXV-5.5).

Montrer qu'une extension séparable est $\{2,3\}$-décomposable si, et seulement si, elle est contenue dans une extension galoisienne dont le groupe de Galois est d'ordre $2^m 3^n$ pour certains n, $m \geqslant 0$.

Indication. On pourra utiliser l'exercice XII-5.

E10. Soit K un corps.

a. Montrer que l'intersection et le compositum d'un nombre fini d'extensions abéliennes de K est encore abélienne. Ces résultats sont-ils encore vrais pour les extensions cycliques ?

b. Soit L/K une extension de degré fini. Montrer l'existence d'une sous-extension abélienne de L/K maximale pour l'inclusion.

E11. Dans ce qui suit, K est un corps **parfait**.

Soit $n \geqslant 2$ et soit $a \in K^\times$. Le but de cet exercice est d'étudier le corps des racines du polynôme $X^n - a \in K[X]$.

On suppose connus les résultats de l'exercice XXVII-9. Jusqu'à la fin, on suppose que le polynôme $X^n - a$ est irréductible. On note L/K le corps des racines de $X^n - a$, et l'on fixe une racine $\alpha \in \overline{K}$ de $X^n - a$.

a. Vérifier que les hypothèses entraînent que n est premier à $\mathrm{car}(K)$.

On pose $E = K(\zeta_n)$, où $\zeta_n \in \overline{K}$ est une racine primitive n-ième de l'unité.

b. Vérifier que $L = E(\alpha)$ et que L/E est cyclique.

c. On suppose que $\mu_n(\overline{K}) \subset K$. Montrer que L/K est abélienne, donner son degré et le treillis de ses sous-extensions.

On suppose que K ne contient pas $\mu_n(\overline{K})$. On désigne par m l'ordre de $\overline{a} \in E^\times/E^{\times n}$.

d. Montrer que $a \in E^{\times \frac{n}{m}}$, et qu'il existe $\beta \in E$ tel que

$$\beta^{\frac{n}{m}} = a \quad \text{et} \quad E \cap K(\alpha) = K(\beta).$$

e. En déduire une factorisation de $X^n - a$ en produit de facteurs irréductibles dans $E[X]$.

f. Montrer que les extensions $E/K(\beta)$ et $K(\alpha)/K(\beta)$ sont linéairement disjointes.

g. Montrer que E ne peut contenir $K(\alpha)$.

h. Montrer que $\mathrm{Gal}(L/E)$ est un sous-groupe distingué de $\mathrm{Gal}(L/K)$. Montrer que $\mathrm{Gal}(L/K) = \mathrm{Gal}(L/E) \rtimes \mathrm{Gal}\big(L/K(\alpha)\big)$ si, et seulement si, E/K et $K(\alpha)/K$ sont linéairement disjointes. En déduire dans ce cas le treillis des sous-extensions de $K(\alpha)/K$.

i. Montrer que la plus grande sous-extension abélienne de L/K contient E. En déduire la plus grande sous-extension abélienne de $K(\alpha)/K$.

j. On suppose maintenant que $n = p^r$, où p est premier et $r \geqslant 1$. Montrer que E/K est cyclique. Pour $\rho \in \mathrm{Gal}(L/K)$, on note

$$\rho(\zeta_{p^r}) = \zeta_{p^r}^{a_\rho}.$$

Montrer que l'application

$$\begin{aligned} \mathrm{Gal}(L/K) &\longrightarrow (\mathbb{Z}/p^r\mathbb{Z})^\times \\ \rho &\longmapsto \overline{a}_\rho \end{aligned}$$

est un morphisme de groupes bien défini.

On note σ un générateur de $\mathrm{Gal}(L/E)$. Soit $\tau \in \mathrm{Gal}(L/K)$ une extension arbitraire d'un générateur de $\mathrm{Gal}(E/K)$. On désigne par p^ℓ l'ordre de $\overline{a} \in K^\times/K^{\times p^r}$, et l'on pose

$$\tau(\alpha) = \zeta_\tau \alpha, \quad \text{où } \zeta_\tau \in \mu_{p^r}(\overline{K}).$$

k. Montrer que $\mathrm{Gal}(L/K) = \langle \sigma, \tau \rangle$.

l. Montrer que a_τ est d'ordre $[E : K]$ modulo p^r, et que $v_p(a_\tau^{[E:K]} - 1) = r$.

m. Montrer que K contient $\mu_{p^i}(\overline{K})$ si, et seulement si, $p^i \mid a_\tau - 1$. Montrer que si $t = v_p(a_\tau - 1) > 0$, alors $[E : K] = p^{r-t}$.

n. Calculer $\tau^{[E:K]}(\alpha)$. Donner une condition nécessaire et suffisante pour que $E(\alpha)/K$ soit cyclique.

o. Décrire le groupe de Galois du corps des racines de $X^5 - 2$ sur $\mathbb{Q}$, puis sur $\mathbb{F}_{11}$.

p. Donner le treillis des sous-extensions de $\mathbb{Q}(\zeta_5)/\mathbb{Q}$, puis étudier le groupe de Galois de $X^{10} - 5 \in \mathbb{Q}[X]$. Donner alors le treillis des sous-extensions de $\mathbb{Q}(\alpha)/\mathbb{Q}$, où $\alpha \in \mathbb{C}$ vérifie $\alpha^{10} = 5$.

E12. Soit L/K une extension galoisienne. On suppose qu'il existe deux sous-extensions F/K et E/K vérifiant les conditions suivantes :

(1) $EF = L$;

(2) $E \cap F = K$;

(3) F/K est galoisienne.

Montrer que $\mathrm{Gal}(L/K) = \mathrm{Gal}(L/F) \rtimes \mathrm{Gal}(L/E)$.

E13. Soit K un corps de caractéristique 0, et soit

$$P = X^3 + pX + q \in K[X]$$

un polynôme irréductible. Enfin, soit $\Delta = \mathrm{disc}(P) = -4p^3 - 27q^2$.

(i) Montrer que $\mathrm{Dec}_K(P) = K(\sqrt{\Delta}, \alpha)$, où $\alpha \in \overline{K}$ est une racine arbitraire de P.

Indication. Soient $\alpha, \beta, \gamma \in \overline{K}$ les racines de P. On pose

$$\delta = (\beta - \alpha)(\gamma - \alpha)(\gamma - \beta).$$

Utiliser les relations coefficients-racines pour montrer successivement que $(\beta - \alpha)(\gamma - \alpha) \in K(\alpha)$, puis que $\beta \in K(\delta, \alpha)$, et enfin l'égalité

$$\mathrm{Dec}_K(P) = K(\delta, \alpha).$$

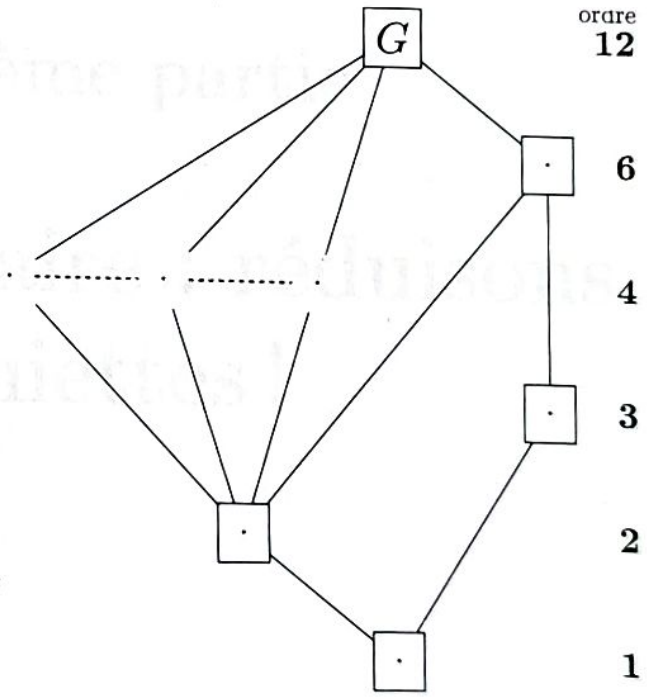
(ii) Montrer que $\mathrm{Gal}_K(P) \simeq \mathbb{Z}/3\mathbb{Z}$ si $\Delta \in K^{\times 2}$, et $\mathrm{Gal}_K(P) \simeq \mathfrak{S}_3$ sinon, et décrire les éléments de $\mathrm{Gal}_K(P)$ dans chaque cas.

E14. Soit $G = \mathbb{Z}/3\mathbb{Z} \times_\rho \mathbb{Z}/4\mathbb{Z}$, où $\rho : \mathbb{Z}/4\mathbb{Z} \longrightarrow \mathrm{Aut}(\mathbb{Z}/3\mathbb{Z})$ est l'unique morphisme de groupes tel que $\rho(\overline{1}) = -\mathrm{Id}_{\mathbb{Z}/3\mathbb{Z}}$.

a. Montrer qu'il existe deux éléments dans G, l'un σ d'ordre 3 et l'autre τ d'ordre 4 tels que

$$\tau\sigma\tau^{-1} = \sigma^{-1}.$$

Montrer alors que G est engendré par σ et τ.
Vérifier que le groupe G possède exactement huit sous-groupes, dont cinq sont distingués et que le treillis de ses sous-groupes est de la forme indiquée ci-contre.
Remplir ce treillis.



Dans la suite, on suppose connus les résultats de l'exercice 13.

b. Soit $P = X^3 - 2X + 1 \in \mathbb{Q}[X]$.

(i) Vérifier que P est irréductible, et admet trois racines réelles α, β, γ, avec $\alpha < \beta < \gamma$.

(ii) Justifier que $(\gamma - \beta)(\gamma - \alpha)(\beta - \alpha) = \sqrt{5}$.

(iii) Justifier que $\mathbb{Q}(\sqrt{5}, \alpha)/\mathbb{Q}$ est galoisienne, de groupe de Galois isomorphe à $\mathfrak{S}_3$, et en expliciter les éléments.

On pose $L = \mathbb{Q}(\zeta_5, \alpha)$.

c. Montrer que $L/\mathbb{Q}$ est galoisienne, de groupe de Galois isomorphe à G.

Indication. Utiliser l'exercice 12.

d. Expliciter deux éléments σ, $\tau \in \mathrm{Gal}(L/\mathbb{Q})$, où σ est d'ordre 3, τ est d'ordre 4, et $\tau\sigma\tau^{-1} = \sigma^{-1}$.

e. Décrire le treillis des sous-extensions de $L/\mathbb{Q}$.

E15. Soit $P \in \mathbb{Q}[X]$ un polynôme irréductible possédant au moins une racine réelle et une racine complexe non réelle. Montrer que $\mathrm{Gal}_{\mathbb{Q}}(P)$ n'est pas abélien.

Cinquième partie

Algèbre linéaire : réduisons en miettes !

Chapitre XXX

Réduction des endomorphismes

Le problème de réduction d'un endomorphisme u d'un K-espace vectoriel E de dimension finie consiste à trouver une base $\mathbf{e}$ de E telle que $\mathrm{Mat}(u\,;\mathbf{e})$ soit la plus « simple » possible.

Il y a deux grandes motivations qui poussent à s'intéresser à ce problème :

(1) faciliter les calculs pratiques de puissances ou d'exponentielle de matrices, qui interviennent dans diverses situations (résolution d'équations différentielles, étude de systèmes dynamiques discrets, etc.) ;

(2) classifier les endomorphismes de E à similitude près.
Matriciellement, le problème de classification s'exprime comme suit : étant données deux matrices A, $A' \in \mathrm{M}_n(K)$, existe-t-il une matrice inversible $P \in \mathrm{GL}_n(K)$ telle que $A = PA'P^{-1}$? Si une telle matrice existe, comment la calculer explicitement ?

Évidemment, le but ultime est de trouver une méthode systématique pour répondre à cette question, et d'obtenir des conditions nécessaires et suffisantes pour que la réponse soit affirmative.

Dans ce chapitre et le suivant, nous étudierons ces deux questions.

1. Polynômes d'endomorphismes

Soit E un K-espace vectoriel, et soit $u \in \mathscr{L}(E)$.

Si $P = a_n X^n + \cdots + a_1 X + a_0 \in K[X]$, on pose

$$P(u) = a_n u^n + \cdots + a_1 u + a_0 \mathrm{Id}_E \in \mathscr{L}(E)$$

où $u^n = u \circ \cdots \circ u$ (avec $u^0 = \mathrm{Id}_E$).

1.1. Définition. Un endomorphisme de E de la forme $P(u)$, avec $P \in K[X]$, est appelé un *polynôme d'endomorphisme en u.*

On note
$$K[u] = \{P(u) \mid P \in K[X]\}.$$

Le lemme suivant est facile, mais fondamental.

1.2. Lemme. *Soit E un K-espace vectoriel, et soit $u \in \mathscr{L}(E)$. Alors, l'application*
$$\begin{aligned} ev_u : K[X] &\longrightarrow \mathscr{L}(E) \\ P &\longmapsto P(u) \end{aligned}$$
est un morphisme de K-algèbres associatives unitaires, et $K[u]$ est une sous-algèbre commutative de $\mathscr{L}(E)$. Autrement dit, pour tous $P, Q \in K[X]$ et tout $\lambda \in K$, on a

(1) $1(u) = \mathrm{Id}_E$;

(2) $(\lambda P)(u) = \lambda P(u)$;

(3) $(P+Q)(u) = P(u) + Q(u)$;

(4) $(PQ)(u) = P(u) \circ Q(u) = Q(u) \circ P(u)$.

De plus, pour tout sous-espace F de E stable par u, on a
$$P(u_F) = P(u)_F, \quad \text{pour tout } P \in K[X].$$

Démonstration. Le fait que ev_u soit K-linéaire provient des définitions de la somme d'un polynôme et de l'évaluation en u.

Soient $P, Q \in K[X]$. Écrivons
$$P = \sum_{\ell \geqslant 0} a_\ell X^\ell, \quad Q = \sum_{m \geqslant 0} b_m X^m,$$
où les scalaires a_ℓ et b_m sont presque tous nuls. Par définition du produit de $K[X]$, on a
$$(PQ)(u) = \sum_{n \geqslant 0} \Big(\sum_{k=0}^{n} a_k b_{n-k} \Big) u^n.$$
Par ailleurs, on a
$$P(u) \circ Q(u) = \Big(\sum_{\ell \geqslant 0} a_\ell u^\ell \Big) \circ \Big(\sum_{m \geqslant 0} b_m u^m \Big).$$
Ainsi, on a
$$P(u) \circ Q(u) = \sum_{\ell, m \geqslant 0} (a_\ell u^\ell) \circ (b_m u^m).$$
La linéarité de u^ℓ montre que
$$(a_\ell u^\ell) \circ (b_m u^m) = a_\ell b_m (u^\ell \circ u^m), \quad \text{pour tous } \ell, m \geqslant 0.$$

En regroupant les termes, on a donc

$$P(u) \circ Q(u) = \sum_{n \geqslant 0} \Big(\sum_{k=0}^{n} a_k b_{n-k} \Big) u^n = (PQ)(u).$$

Enfin, on a $(PQ)(u) = (QP)(u)$, et donc $P(u) \circ Q(u) = Q(u) \circ P(u)$. Le fait que $K[u]$ soit une sous-algèbre commutative de $\mathscr{L}(u)$ provient alors des points (1) – (4).

La dernière partie du lemme découle des définitions. Cela achève la démonstration. □

Les polynômes d'endomorphismes permettent d'obtenir à peu de frais des sous-espaces stables.

1.3. Exemple. Soit $u \in \mathscr{L}(E)$ et soit $S \in K[X]$. Alors, $\operatorname{Ker}\big(S(u)\big)$ est un sous-espace vectoriel de E stable par u.

En effet, si $x \in \operatorname{Ker}\big(S(u)\big)$, alors on a

$$S(u)\big(u(x)\big) = (S(u) \circ u)(x) = \big(u \circ S(u)\big)(x) = u\big(S(u)(x)\big) = u(0) = 0,$$

et donc $u(x) \in \operatorname{Ker}\big(S(u)\big)$.

Nous avons déjà souligné l'importance du rôle des décompositions en somme directe de sous-espaces stables dans le problème de la réduction des endomorphismes (cf. remarque III-4.9). Les polynômes d'endomorphismes permettent souvent d'obtenir de telles décompositions, grâce au lemme fondamental suivant.

1.4. Lemme (de décomposition des noyaux). *Soient $P_1, \ldots, P_r$ des polynômes de $K[X]$ premiers entre eux deux à deux. Pour tout $u \in \mathscr{L}(E)$, on a*

$$\operatorname{Ker}\big((P_1 \cdots P_r)(u)\big) = \operatorname{Ker}\big(P_1(u)\big) \oplus \cdots \oplus \operatorname{Ker}\big(P_r(u)\big).$$

Démonstration. Puisque les polynômes $P_1 \cdots P_{r-1}$ et P_r sont premiers entre eux, il suffit de traiter le cas $r = 2$, et de finir par récurrence. Supposons donc $r = 2$, et posons $P = P_1 P_2$. Montrons tout d'abord que

$$\operatorname{Ker}\big(P(u)\big) = \operatorname{Ker}\big(P_1(u)\big) + \operatorname{Ker}\big(P_2(u)\big).$$

Rappelons que l'on a

$$P(u) = P_1(u) \circ P_2(u) = P_2(u) \circ P_1(u).$$

Si $x_i \in \mathrm{Ker}\big(P_i(u)\big)$, on a

$$\begin{aligned} P(u)(x_1+x_2) &= P(u)(x_1)+P(u)(x_2) \\ &= P_2(u)\big(P_1(u)(x_1)\big)+P_1(u)\big(P_2(u)(x_2)\big) \\ &= P_2(u)(0)+P_1(u)(0) \\ &= 0. \end{aligned}$$

Ainsi, on a $\mathrm{Ker}\big(P_1(u)\big)+\mathrm{Ker}\big(P_2(u)\big) \subset \mathrm{Ker}\big(P(u)\big)$.

Inversement, soit $x \in \mathrm{Ker}\big(P(u)\big)$. Puisque P_1 et P_2 sont premiers entre eux, il existe U, $V \in K[X]$ tels que

$$UP_1+VP_2=1.$$

On a alors $\mathrm{Id}_E=(UP_1)(u)+(VP_2)(u)$. Posons

$$x_1=(VP_2)(u)(x), \quad x_2=(UP_1)(u)(x).$$

On a donc $x=x_1+x_2$. Vérifions que $x_i \in \mathrm{Ker}\big(P_i(u)\big)$. On a

$$P_1(u)(x_1)=\big(P_1(u)\circ(VP_2)(u)\big)(x)=(P_1VP_2)(u)(x).$$

On a donc

$$P_1(u)(x_1)=(VP)(u)(x)=V(u)\big(P(u)(x)\big)=V(u)(0)=0.$$

Ainsi, $x_1 \in \mathrm{Ker}\big(P_1(u)\big)$. On montre de même que $x_2 \in \mathrm{Ker}\big(P_2(u)\big)$. On a par conséquent l'égalité

$$\mathrm{Ker}\big(P(u)\big)=\mathrm{Ker}\big(P_1(u)\big)+\mathrm{Ker}\big(P_2(u)\big).$$

Il reste à montrer que la somme est directe.

Soit $x \in \mathrm{Ker}\big(P_1(u)\big) \cap \mathrm{Ker}\big(P_2(u)\big)$. On a alors

$$\begin{aligned} x &= (UP_1)(u)(x)+(VP_2)(u)(x) \\ &= U(u)\big(P_1(u)(x)\big)+V(u)\big(P_2(u)(x)\big) \\ &= U(u)(0)+V(u)(0) \\ &= 0. \end{aligned}$$

Cela montre que la somme est directe, et achève la démonstration. □

Au vu de l'exemple 1.3, le lemme de décomposition des noyaux est particulièrement intéressant lorsque $P(u)=0$, car on obtient alors une décomposition de E en somme directe de sous-espaces stables par u. Nous allons maintenant étudier l'existence de tels polynômes.

1.5. Définition. Soit $u \in \mathscr{L}(E)$. *L'annulateur* de u, noté $\mathrm{Ann}(u)$ est l'ensemble

$$\mathrm{Ann}(u)=\{P \in K[X] \mid P(u)=0\}.$$

Un élément de $\mathrm{Ann}(u)$ est appelé un *polynôme annulateur* de u.

Attention ! L'annulateur de u peut être nul (cf. exercice 1).

1.6. Lemme. *Soit $u \in \mathscr{L}(E)$. Alors, $\mathrm{Ann}(u)$ est un idéal de $K[X]$.*

Démonstration. Par définition, $\mathrm{Ann}(u)$ est le noyau de ev_u, qui est un morphisme de K-algèbres associatives unitaires. Or, un tel morphisme est un morphisme d'anneaux, d'où le résultat par le lemme XV-4.5. □

1.7. Proposition. *Soit E un K-espace vectoriel de dimension finie, et soit $u \in \mathscr{L}(E)$. Alors, $\mathrm{Ann}(u)$ est non nul. De plus, il existe un unique polynôme unitaire $\mu_u \in K[X]$ tel que*

$$\mathrm{Ann}(u) = \mu_u K[X] = \{\mu_u Q \mid Q \in K[X]\}.$$

Démonstration. Comme E est de dimension finie, disons n, l'espace $\mathscr{L}(E)$ est de dimension n^2. Alors, la famille $\left(\mathrm{Id}_E, u, \ldots, u^{n^2}\right)$ est de cardinal n^2+1, donc nécessairement liée. Il existe donc $a_0, \ldots, a_{n^2} \in K$ **non tous nuls** tels que

$$a_0 \mathrm{Id}_E + a_1 u + \cdots + a_{n^2} u^{n^2} = 0.$$

Ainsi, $P = a_{n^2} X^{n^2} + \cdots + a_1 X + a_0$ est un élément non nul de $\mathrm{Ann}(u)$. Comme $K[X]$ est principal par le lemme XV-6.14, l'idéal non nul $\mathrm{Ann}(u)$ est engendré par un unique polynôme unitaire μ_u (cf. par exemple, la remarque XVIII-5.11). Cela achève la démonstration. □

On supposera dorénavant que E est de dimension finie jusqu'à la fin de ce chapitre, sauf mention explicite du contraire.

1.8. Définition. Le polynôme μ_u est appelé le *polynôme minimal* de u. Il est caractérisé de manière unique par les propriétés suivantes :

(1) le polynôme μ_u est unitaire ;

(2) $\mu_u(u) = 0$;

(3) pour tout $P \in K[X]$, on a $P(u) = 0 \implies \mu_u \mid P$.

De manière plus condensée, μ_u est l'unique polynôme unitaire de $K[X]$ tel que

$$\text{pour tout } P \in K[X],\ P(u) = 0 \iff \mu_u \mid P.$$

1.9. Remarque. Il est à noter que μ_u n'est pas constant, sauf si $E = \{0\}$. En effet, si μ_u est constant, alors $\mu_u = 1$, et donc $\mu_u(u) = 0 = \mathrm{Id}_E$, ce qui implique que $E = \{0\}$.

Pour calculer μ_u, il « suffit » de trouver un polynôme annulateur P de degré assez bas si possible, puis de faire la liste de tous les diviseurs unitaires de P. Le polynôme μ_u sera alors le diviseur unitaire annulant u de plus bas degré. La démonstration de la proposition précédente fournit (tout du moins en théorie) un polynôme annulateur de u de degré n^2, où $n = \dim_K(E)$, ce qui est un peu grand. Le but de ce qui suit est de définir un polynôme annulateur de u de degré n.

Notons tout d'abord que si $M \in \mathrm{M}_n(K)$, la matrice $X\mathrm{I}_n - M$ est à coefficients dans $K[X]$, et ainsi $\det(X\mathrm{I}_n - M) \in K[X]$. On a alors la définition suivante.

1.10. Définition. Le *polynôme caractéristique* d'une matrice $M \in \mathrm{M}_n(K)$ est le polynôme $\chi_M \in K[X]$ défini par

$$\chi_M = \det(X\mathrm{I}_n - M).$$

1.11. Lemme. *Si $M \in \mathrm{M}_n(K)$, le polynôme χ_M est unitaire de degré n. Si, de plus, on écrit $\chi_M = X^n + a_{n-1}X^{n-1} + \cdots + a_0$, alors on a*

$$\mathrm{tr}(M) = -a_{n-1} \quad \text{et} \quad \det(M) = (-1)^n a_0.$$

Démonstration. Si $M = (a_{ij})_{i,j}$, on a $X\mathrm{I}_n - M = (X\delta_{ij} - a_{ij})_{i,j}$, et donc

$$\chi_M = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma)(X\delta_{\sigma(1)1} - a_{\sigma(1)1}) \cdots (X\delta_{\sigma(n)n} - a_{\sigma(n)n}).$$

Si $n = 1$, le résultat est clair. On peut donc supposer que $n \geqslant 2$. Chaque terme du produit dans le terme général de la somme est de degré $\leqslant 1$, et donc le terme général est de degré $\leqslant n$. De plus, si $\sigma \neq \mathrm{Id}$, alors σ possède au moins deux points qui ne sont pas fixes. En effet, si σ fixe $n-1$ points, il fixe également le dernier, puisque σ est bijective. Ainsi, si $\sigma \neq \mathrm{Id}$, il existe deux entiers $i, j \in [\![1, n]\!]$ avec $i \neq j$ tels que $\sigma(i) \neq i$ et $\sigma(j) \neq j$. Dans ce cas, on a $X\delta_{\sigma(i)i} - a_{\sigma(i)i} = -a_{\sigma(i)i}$, qui est soit nul, soit de degré 0, et de même pour $X\delta_{\sigma(j)j} - a_{\sigma(j)j}$. Mais alors, le terme correspondant à σ est de degré inférieur à $n-2$.

Le seul terme incluant éventuellement des termes en X^n et X^{n-1} est donc celui qui est obtenu pour $\sigma = \mathrm{Id}$, à savoir

$$(X - a_{11}) \cdots (X - a_{nn}).$$

Ce polynôme est unitaire de degré n, et son coefficient en X^{n-1} est

$$-(a_{11} + \cdots + a_{nn}) = -\mathrm{tr}(M).$$

Enfin, le terme constant de χ_M est clairement $\det(-M) = (-1)^n \det(M)$. On en déduit le résultat voulu. □

1.12. Remarque. La proposition IV-4.8 donne immédiatement le résultat suivant : si $M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$, on a

$$\chi_M = \chi_A \chi_C.$$

Cette propriété sera utilisée constamment dans la suite.

On continue par un exemple de calcul de polynôme caractéristique.

1.13. Exemple. Soit $M \in \mathrm{M}_n(K)$ la matrice définie par

$$M = \begin{pmatrix} 0 & 0 & & 0 & -a_0 \\ 1 & 0 & & & -a_1 \\ 0 & 1 & & & -a_2 \\ & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}.$$

Alors, $\chi_M = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0$.

En effet, on a

$$\chi_M = \begin{vmatrix} X & 0 & & & a_0 \\ -1 & X & & & a_1 \\ 0 & -1 & & & a_2 \\ & & \ddots & & \vdots \\ 0 & 0 & \cdots & -1 & X + a_{n-1} \end{vmatrix}.$$

Posons $f = X^n + a_{n-1}X^{n-1} + \cdots + a_0$. En faisant l'opération

$$L_1 \longleftrightarrow L_1 + XL_2 + \cdots + X^{n-1}L_n,$$

on obtient

$$\chi_M = \begin{vmatrix} 0 & 0 & & & f \\ -1 & X & & & a_1 \\ 0 & -1 & & & a_2 \\ & & \ddots & & \vdots \\ 0 & 0 & \cdots & -1 & X + a_{n-1} \end{vmatrix}.$$

En développant par rapport à la première ligne, on obtient

$$\chi_M = (-1)^{1+n} f \begin{vmatrix} -1 & * & * & * \\ & -1 & * & * \\ & & \ddots & * \\ & & & -1 \end{vmatrix} = \big((-1)^{n+1} f\big)(-1)^{n-1} = f.$$

Revenons à des considérations plus générales.

Soit $u \in \mathscr{L}(E)$, et soient $\mathbf{e}$, $\mathbf{e}'$ deux bases de E. Soient $M = \mathrm{Mat}(u;\mathbf{e})$ et $M' = \mathrm{Mat}(u;\mathbf{e}')$. Il existe alors une matrice[1] P dans $\mathrm{GL}_n(K)$ telle que $M = PM'P^{-1}$. On a alors

$$\begin{aligned}\chi_M &= \det(X\mathrm{I}_n - M)\\ &= \det(X\mathrm{I}_n - PM'P^{-1})\\ &= \det\left(P(X\mathrm{I}_n - M')P^{-1}\right)\\ &= \det(P)\det(X\mathrm{I}_n - M')\det(P)^{-1}\\ &= \chi_{M'}.\end{aligned}$$

Ainsi, le polynôme caractéristique d'une matrice représentative de u ne dépend pas de la base choisie.

1.14. Définition. Soit $u \in \mathscr{L}(E)$. Le *polynôme caractéristique* de u est le polynôme caractéristique d'une matrice représentative de u dans une base arbitraire. On le note χ_u.

Une retraduction du lemme 1.11 en termes d'endomorphismes donne immédiatement ce qui suit.

1.15. Lemme. *Soit E un K-espace vectoriel de dimension finie $n \geqslant 1$, et soit $u \in \mathscr{L}(E)$. Alors, χ_u est un polynôme unitaire de degré n. De plus, si l'on écrit $\chi_u = X^n + a_{n-1}X^{n-1} + \cdots + a_0$, on a les égalités*

$$\mathrm{tr}(u) = -a_{n-1} \quad \textit{et} \quad \det(u) = (-1)^n a_0.$$

On a alors le résultat suivant.

1.16. Théorème (Cayley-Hamilton). *Pour tout $u \in \mathscr{L}(E)$, le polynôme caractéristique χ_u est un polynôme annulateur de u, i.e. $\chi_u(u) = 0$.*

Autrement dit, $\mu_u \mid \chi_u$.

Démonstration. Soit $n = \dim_K(E)$.

Commençons par établir le théorème dans un cas particulier. On suppose qu'il existe $v \in E$ tel que $\big(v, u(v), \ldots, u^{n-1}(v)\big)$ soit une base de E. En particulier, il existe $a_0, \ldots, a_{n-1} \in K$ tels que

$$u^n(v) = -a_0 v - a_1 u(v) - \cdots - a_{n-1}u^{n-1}(v).$$

1. À savoir, la matrice de passage de $\mathbf{e}'$ à $\mathbf{e}$.

La matrice représentative de u dans cette base est alors

$$\begin{pmatrix} 0 & 0 & & 0 & -a_0 \\ 1 & 0 & & & -a_1 \\ 0 & 1 & & & -a_2 \\ & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}.$$

D'après l'exemple précédent, on a donc

$$\chi_u = X^n + a_{n-1}X^{n-1} + \cdots + a_0.$$

Pour montrer que $\chi_u(u) = 0$, il suffit de montrer que $\chi_u(u)\big(u^m(v)\big) = 0$ pour $m \in [\![0, n-1]\!]$, puisque $\big(v, u(v), \ldots, u^{n-1}(v)\big)$ est une base de E. Mais on a

$$\chi_u(u)\big(u^m(v)\big) = \big(\chi_u(u) \circ u^m\big)(v) = \big(u^m \circ \chi_u(u)\big)(v) = u^m\big(\chi_u(u)(v)\big).$$

Il suffit donc de montrer que $\chi_u(u)(v) = 0$. Or,

$$\chi_u(u)(v) = u^n(v) + a_{n-1}u^{n-1}(v) + \cdots + a_1u(v) + a_0v = 0,$$

par définition des a_i. Le théorème est donc démontré dans ce cas particulier.

L'idée est de se ramener à ce cas particulier.

Pour montrer que $\chi_u(u) \in \mathscr{L}(E)$ est l'endomorphisme nul, il faut montrer que $\chi_u(u)(v) = 0$ pour tout $v \in E$. Clairement, on a $\chi_u(u)(0) = 0$. Soit maintenant $v \in E \setminus \{0\}$. Puisque E est de dimension $n \geqslant 1$, il existe un entier $d \in [\![1, n]\!]$ tel que la famille $\big(v, u(v), \ldots, u^{d-1}(v)\big)$ soit libre et de cardinal maximal. Alors, la famille $\big(v, u(v), \ldots, u^d(v)\big)$ est liée, et donc il existe $\lambda_0, \ldots, \lambda_d \in K$ non tous nuls tels que

$$\lambda_0 v + \cdots + \lambda_d u^d(v) = 0.$$

On a $\lambda_d \neq 0$, sinon tous les λ_i seraient nuls, la famille $\big(v, u(v), \ldots, u^{d-1}(v)\big)$ étant libre. Ainsi, la relation précédente s'écrit

$$u^d(v) = -a_0v - a_1u(v) - \cdots - a_{d-1}u^{d-1}(v), a_i \in K.$$

En particulier, le sous-espace $F = \mathrm{Vect}_K\big(v, u(v), \ldots, u^{d-1}(v)\big)$ est stable par u, et $\mathbf{f} = \big(v, u(v), \ldots, u^{d-1}(v)\big)$ en est une base. D'après le premier point, on a

$$\chi_{u_F}(u_F) = 0.$$

Complétons $\big(v, u(v), \ldots, u^{d-1}(v)\big)$ en une base $\mathbf{e}$ de E. Soit $M = \mathrm{Mat}(u\,;\mathbf{e})$. Alors, M est une matrice triangulaire supérieure par blocs de la forme

$$M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix},$$

où $A = \mathrm{Mat}(u_F\,;\mathbf{f})$.

On a alors donc $\chi_M = \chi_A \chi_C$, et donc

$$\chi_u(u) = \chi_M(u) = \chi_A(u) \circ \chi_C(u) = \chi_C(u) \circ \chi_A(u).$$

Pour montrer que $\chi_u(u)(v) = 0$, il suffit donc de voir que $\chi_A(u)(v) = 0$, puisque $\chi_C(u)(0) = 0$. Or, étant donné que $v \in F$, on a

$$\chi_A(u)(v) = \chi_A(u)_F(v) = \chi_A(u_F)(v) = \chi_{u_F}(u_F)(v) = 0.$$

Cela achève la démonstration. □

Toutes les notions et résultats précédents ont leurs équivalents matriciels. Si $M \in \mathrm{M}_n(K)$, et $P = a_d X^d + \cdots + a_1 X + a_0 \in K[X]$, on pose

$$P(M) = a_d M^d + \cdots + a_1 M + a_0 \mathrm{I}_n \in \mathrm{M}_n(K).$$

Il est facile de constater que, pour tout $u \in \mathscr{L}(E)$, toute base $\mathbf{e}$ de E, et tout $P \in K[X]$, on a

$$\mathrm{Mat}\big(P(u)\,;\mathbf{e}\big) = P\big(\mathrm{Mat}(u\,;\mathbf{e})\big).$$

En appliquant tout ce qui précède à l'endomorphisme

$$\begin{aligned} K^n &\longrightarrow K^n \\ X &\longmapsto MX, \end{aligned}$$

on obtient le résultat suivant.

1.17. Théorème. *Soit $M \in \mathrm{M}_n(K)$. Alors, pour tous $P, Q \in K[X]$, et tout $\lambda \in K$, on a*

(1) $0(M) = 0$, *et* $1(M) = \mathrm{I}_n$;

(2) $(P+Q)(M) = P(M) + Q(M)$;

(3) $(PQ)(M) = P(M)Q(M) = Q(M)P(M)$;

(4) $(\lambda P)(M) = \lambda P(M)$.

De plus, l'ensemble $\mathrm{Ann}(M) = \{P \mid P(M) = 0\}$ *des polynômes annulateurs de M est non nul, et il existe un unique polynôme unitaire $\mu_M \in K[X]$ tel que*

$$\mathrm{Ann}(M) = \mu_M K[X].$$

Enfin, le polynôme caractéristique χ_M vérifie $\chi_M(M) = 0$. Autrement dit, on a $\mu_M \mid \chi_M$.

1.18. Remarque. Pour tout $u \in \mathscr{L}(E)$, toute base $\mathbf{e}$ de E, et tout polynôme $P \in K[X]$, on a

$$P(u) = 0 \iff P\big(\mathrm{Mat}(u;\mathbf{e})\big) = 0.$$

On a donc $\mathrm{Ann}(u) = \mathrm{Ann}\big(\mathrm{Mat}(u;\mathbf{e})\big)$, et par suite

$$\mu_u = \mu_{\mathrm{Mat}(u;\mathbf{e})}.$$

1.19. Proposition. *Soit E un K-espace vectoriel de dimension $n \geqslant 1$, et soit $u \in \mathscr{L}(E)$. Alors, $\mu_u \mid \chi_u \mid \mu_u^n$. En particulier, μ_u et χ_u ont même racines dans K.*

Plus généralement, les polynômes μ_u et χ_u ont les mêmes diviseurs irréductibles dans $K[X]$.

Démonstration. Le fait que $\mu_u \mid \chi_u$ vient du fait que χ_u est un polynôme annulateur de u par le théorème de Cayley-Hamilton. Montrons que $\chi_u \mid \mu_u^n$. Notons que, pour tout $m \geqslant 1$, on a l'égalité

$$X^m - Y^m = (X-Y)(X^{m-1} + X^{m-2}Y + \cdots + XY^{m-2} + Y^{m-1}) \in K(X)[Y].$$

Si l'on note d le degré de μ_u, on en déduit facilement l'existence de d polynômes $P_0, \ldots, P_{d-1} \in K[X]$ tels que

$$\mu_u(X) - \mu_u(Y) = (X-Y)(P_0 + P_1 Y + \cdots + P_{d-1}Y^{d-1}) \in K(X)[Y].$$

Soit M la matrice représentative de u dans une base fixée. On a donc

$$\mu_u(M) = 0 \in \mathrm{M}_n(K) \subset \mathrm{M}_n\big(K(X)\big).$$

En considérant M comme un élément de $\mathrm{M}_n\big(K(X)\big)$, l'égalité précédente implique

$$\mu_u(X)\mathrm{I}_n = (X\mathrm{I}_n - M)Q \in \mathrm{M}_n\big(K(X)\big),$$

où $Q = (P_0\mathrm{I}_n + P_1 M + \cdots + P_{d-1}M^{d-1}) \in \mathrm{M}_n\big(K(X)\big)$, puisque $\mu_M(M) = 0$. En prenant le déterminant des deux côtés, on obtient

$$\mu_u^n = \chi_u \det(Q) \in K(X).$$

Remarquons maintenant que $Q \in \mathrm{M}_n(K[X])$. On a alors $\det(Q) \in K[X]$, d'où le résultat voulu.

Soit $\pi \in K[X]$ un polynôme irréductible. Si $\pi \mid \mu_u$, alors $\pi \mid \chi_u$, car $\mu_u \mid \chi_u$. Inversement, si $\pi \mid \chi_u$, alors $\pi \mid \mu_u^n$, et donc $\pi \mid \mu_u$ par le lemme d'Euclide. Pour obtenir l'assertion manquante, il suffit d'appliquer ce qui précède à un diviseur irréductible de la forme $X - \lambda$, $\lambda \in K$, puisque $\lambda \in K$ est racine d'un polynôme $P \in K[X]$ si, et seulement si, $X - \lambda$ divise P. Cela achève la démonstration. □

On continue par un résultat qui permet de réduire le calcul du polynôme minimal au calcul de polynômes minimaux de sous-espaces stables de dimension plus petites.

1.20. Lemme. *Soit $u \in \mathscr{L}(E)$, et soit $E = E_1 \oplus \cdots \oplus E_r$ une décomposition de E en somme directe de sous-espaces stables par u.*

Alors, on a

$$\mu_u = \operatorname{ppcm}(\mu_{u_{E_1}}, \ldots, \mu_{u_{E_r}}).$$

Démonstration. Pour tout $P \in K[X]$, et tout $x_i \in E_i$, on a

$$\begin{aligned} P(u)(x_1 + \cdots + x_r) &= P(u)(x_1) + \cdots + P(u)(x_r) \\ &= P(u_{E_1})(x_1) + \cdots P(u_{E_r})(x_r). \end{aligned}$$

Comme $E_1, \ldots, E_r$ sont en somme directe, on a

$$P(u)(x_1 + \cdots + x_r) = 0 \iff P(u_{E_i})(x_i) = 0, \text{ pour tout } i \in [\![1, r]\!].$$

Comme tout $x \in E$ s'écrit de manière unique sous la forme

$$x = x_1 + \cdots + x_r, \quad x_i \in E_i,$$

on obtient aisément

$$P(u) = 0 \iff P(u_{E_i}) = 0, \text{ pour tout } i \in [\![1, r]\!].$$

Autrement dit, pour tout $P \in K[X]$, on a $P(u) = 0$ si, et seulement si, $\mu_{u_{E_i}} \mid P$ pour tout $i \in [\![1, r]\!]$. Ainsi, $P(u) = 0$ si, et seulement si, on a $\operatorname{ppcm}(\mu_{uE_1}, \ldots, \mu_{u_{E_r}}) \mid P$. Le polynôme $\operatorname{ppcm}(\mu_{u_{E_1}}, \ldots, \mu_{u_{E_r}})$ étant unitaire, on en déduit le résultat. □

1.21. Corollaire. *Soit $u \in \mathscr{L}(E)$. Si $\mu_u = P_1 \cdots P_r$, où $P_1, \ldots, P_r \in K[X]$ sont unitaires et premiers entre eux deux à deux, alors*

$$\mu_{u_{\operatorname{Ker}(P_i(u))}} = P_i, \quad \textit{pour tout } i \in [\![1, r]\!].$$

Démonstration. Le lemme de décomposition des noyaux nous donne

$$E = \operatorname{Ker}\big(\mu_u(u)\big) = \operatorname{Ker}\big(P_1(u)\big) \oplus \cdots \oplus \operatorname{Ker}\big(P_r(u)\big).$$

Pour tout entier $i \in [\![1, r]\!]$, notons μ_i le polynôme minimal de $u_{\operatorname{Ker}(P_i(u))}$. Par définition, P_i annule $u_{\operatorname{Ker}(P_i(u))}$, et donc $\mu_i \mid P_i$. Comme $P_1, \ldots, P_r$ sont premiers entre eux deux à deux , il en est alors de même de $\mu_1, \ldots, \mu_r$. Le lemme 1.20 donne alors

$$\mu_u = \operatorname{ppcm}(\mu_1, \ldots, \mu_r) = \mu_1 \cdots \mu_r.$$

On a donc en particulier

$$\deg(\mu_u) = \deg(P_1) + \cdots + \deg(P_r) = \deg(\mu_1) + \cdots + \deg(\mu_r).$$

Mais $\mu_i \mid P_i$, et donc $\deg(\mu_i) \leqslant \deg(P_i)$, pour tout $i \in [\![1,r]\!]$. On a alors nécessairement $\deg(\mu_i) = \deg(P_i)$ pour tout $i \in [\![1,r]\!]$, et comme μ_i est un diviseur de P_i, on a $\mu_i = P_i$, car ces deux polynômes sont unitaires. Cela achève la démonstration. □

Finissons cette section par un analogue des résultats précédents, mais concernant cette fois le polynôme caractéristique.

1.22. Lemme. *Soit $u \in \mathscr{L}(E)$, et soit $E = E_1 \oplus \cdots \oplus E_r$ une décomposition de E en somme directe de sous-espaces stables par u. Alors, on a*

$$\chi_u = \chi_{u_{E_1}} \cdots \chi_{u_{E_r}}.$$

Démonstration. Pour tout $i \in [\![1,r]\!]$, choisissons une base $\mathbf{e}_i$ de E_i. Si $\mathbf{e}$ est la base de E obtenue en recollant ces bases, alors on a

$$\mathrm{Mat}(u\,;\mathbf{e}) = \begin{pmatrix} \mathrm{Mat}(u_{E_1};\mathbf{e}_1) & & \\ & \ddots & \\ & & \mathrm{Mat}(u_{E_r};\mathbf{e}_r) \end{pmatrix}.$$

Le résultat s'obtient en utilisant les propriétés du déterminant et la définition du polynôme caractéristique d'un endomorphisme. □

1.23. Corollaire. *Soit $u \in \mathscr{L}(E)$. Si $\chi_u = P_1 \cdots P_r$, où $P_1, \ldots, P_r \in K[X]$ sont unitaires et premiers entre eux deux à deux, alors*

$$\chi_{u_{\mathrm{Ker}(P_i(u))}} = P_i, \quad \text{pour tout } i \in [\![1,r]\!].$$

En particulier, $\dim_K\big(\mathrm{Ker}\big(P_i(u)\big)\big) = \deg(P_i)$.

Démonstration. On commence par le cas où $P_i = \pi_i^{m_i}$, où $\pi_i \in K[X]$ est irréductible unitaire. On a donc

$$\chi_u = \pi_1^{m_1} \cdots \pi_r^{m_r}.$$

Notons χ_i le polynôme caractéristique de $u_{\mathrm{Ker}(P_i(u))}$, et μ_i son polynôme minimal. Notons que P_i annule $u_{\mathrm{Ker}(P_i(u))}$, donc $\mu_i \mid P_i$. En particulier, μ_i est une puissance de π_i. Mais alors, χ_i est une puissance de π_i par la proposition 1.19. Écrivons alors $\chi_i = \pi_i^{s_i}$. Le lemme précédent nous donne

$$\chi_u = \chi_1 \cdots \chi_r = \pi_1^{s_1} \cdots \pi_r^{s_r}.$$

Par hypothèse sur $P_1, \ldots, P_r$, les polynômes $\pi_1, \ldots, \pi_r$ sont deux à deux distincts. On en déduit alors que $s_i = m_i$ pour tout $i \in [\![1,r]\!]$, d'où

$$\chi_{u_{\mathrm{Ker}(P_i(u))}} = \pi_i^{m_i} = P_i.$$

Revenons maintenant au cas général, et écrivons

$$P_i = \pi_{i1}^{m_{i1}} \cdots \pi_{ir_i}^{m_{ir_i}},$$

où les $\pi_{i1}, \ldots, \pi_{ir_i}$ sont des polynômes irréductibles unitaires deux à deux distincts. Le lemme de décomposition des noyaux donne

$$\operatorname{Ker}\big(P_i(u)\big) = \operatorname{Ker}\big(\pi_{i1}^{m_{i1}}(u)\big) \oplus \cdots \oplus \operatorname{Ker}\big(\pi_{ir_i}^{m_{ir_i}}(u)\big).$$

On a alors

$$E = \bigoplus_{i=1}^{r} \bigoplus_{j=1}^{r_i} \operatorname{Ker}\big(\pi_{ij}^{m_{ij}}(u)\big).$$

Comme $P_1, \ldots, P_r$ sont premiers deux à deux, les polynômes π_{ij} sont tous distincts, et ainsi les polynômes $\pi_{ij}^{m_{ij}}$ sont deux à deux distincts. Par le point précédent, le polynôme caractéristique de la restriction de u à $\operatorname{Ker}\big(\pi_{ij}^{m_{ij}}(u)\big)$ est $\pi_{ij}^{m_{ij}}$.

Mais alors, le lemme 1.22 montre que le polynôme caractéristique de l'endomorphisme $u_{\operatorname{Ker}(P_i(u))}$ est $\pi_{i1}^{m_{i1}} \cdots \pi_{ir_i}^{m_{ir_i}}$, c'est-à-dire P_i.

Le dernier point est clair. □

2. Diagonalisation

Le but de cette section et des suivantes est de démontrer l'existence de bases de E dans lesquelles la matrice représentative d'un endomorphisme à une forme particulière qui la rend sympathique pour les calculs (sous certaines conditions). En effet, il arrive souvent que l'on doive calculer les puissances successives d'une matrice $A \in \mathrm{M}_n(K)$, par exemple pour ensuite calculer l'exponentielle d'une matrice, (afin de résoudre un système d'équations différentielles linéaires), ou pour étudier le comportement de la suite $(A^m)_{m\geqslant 0}$ lorsque $m \longrightarrow \infty$.

Cette simple question peut être très ardue.

Par exemple, si $A = \begin{pmatrix} 7 & -1 \\ -2 & 8 \end{pmatrix}$, comment calculer A^{2017}, ou même A^m pour m arbitraire[2] ?

Si $A \in \mathrm{M}_n(K)$, notons $u \in \mathscr{L}(K^n)$ l'endomorphisme de K^n canoniquement associé à la matrice A. L'idée est de trouver une décomposition de K^n en somme directe sous-espaces $E_1, \ldots, E_r$ stables par u, et des bases $\mathbf{e}_1, \ldots, \mathbf{e}_r$ de $E_1, \ldots, E_r$ respectivement, de façon à ce que les matrices $\operatorname{Mat}(u_{E_i}; \mathbf{e}_i)$ soient les plus simples possibles.

2. Le lecteur peu convaincu de la difficulté de ce problème pourra essayer de trouver à la main une formule générale pour A^m.

Vu l'exemple précédent, calculer les puissances d'une matrice, même de taille 2×2, n'est pas nécessairement si facile. La situation idéale est lorsque chaque E_i est de dimension 1 (et donc $r = \dim_K K^n = n$). Si $e_i \in E_i$ est non nul, puisque E_i est stable par u, il existe $\lambda_i \in K$ tel que $u(e_i) = \lambda_i e_i$.

Dans ce cas, en posant $\mathbf{e} = (e_1, \ldots, e_n)$, on voit que $\mathrm{Mat}(u\,;\mathbf{e})$ est diagonale. Matriciellement, cela correspond à l'existence d'une matrice $P \in \mathrm{GL}_n(K)$ et une matrice diagonale $D \in \mathrm{M}_n(K)$ telle que $A = PDP^{-1}$.

Pour tout $m \geqslant 0$, on a alors $A^m = PD^mP^{-1}$, ce qui réduit le calcul des puissances de A au calcul des puissances de D, ce qui est extrêmement simple.

Par exemple, on peut vérifier, à l'aide des techniques exposées ci-après, que

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 6 & 0 \\ 0 & 9 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}^{-1}.$$

Pour tout $m \geqslant 0$, on a ainsi

$$A^m = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix} \begin{pmatrix} 6^m & 0 \\ 0 & 9^m \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}^{-1}.$$

Une question naturelle se pose alors : existe-t-il toujours une matrice inversible P et une matrice diagonale D telles que $A = PDP^{-1}$?

La réponse est non en général, comme on peut le voir en prenant par exemple[3] $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$.

En termes d'endomorphismes, cela revient à demander qu'il existe une base dans laquelle la matrice représentative de l'endomorphisme de K^n représenté par A soit diagonale. Commençons par traduire cela de manière plus intrinsèque.

2.1. Lemme. *Soit E un K-espace vectoriel de dimension finie non nulle, soit $u \in \mathscr{L}(E)$ et soit $\mathbf{e} = (e_1, \ldots, e_n)$ une base de E. Les propriétés suivantes sont équivalentes :*

(1) *la matrice $\mathrm{Mat}(u\,;\mathbf{e})$ est diagonale ;*

(2) *il existe $\lambda_1, \ldots, \lambda_n \in K$ tels que $u(e_i) = \lambda_i e_i$, pour tout $i \in [\![1, n]\!]$.*

Démonstration. Cela découle immédiatement de la définition d'une matrice représentative. □

Remarquons au passage que les vecteurs e_i sont non nuls, puisque faisant partie d'une base. Ce lemme justifie alors la définition suivante.

3. La justification de cette affirmation péremptoire est laissée en exercice au lecteur.

2.2. Définition. Soit $u \in \mathscr{L}(E)$. On dit que $\lambda \in K$ est une *valeur propre* de u s'il existe $x \in E \setminus \{0\}$ tel que $u(x) = \lambda x$. Un tel vecteur $x \in E$ est alors appelé un *vecteur propre* de u associé à la valeur propre λ. Cela revient à dire que Kx est une droite stable par u.

Si $\lambda \in K$ est une valeur propre de u, le *sous-espace propre* E_λ associé est le sous-espace vectoriel

$$E_\lambda = \mathrm{Ker}(u - \lambda \mathrm{Id}_E).$$

C'est donc l'ensemble des vecteurs propres associés à la valeur propre λ, auquel on ajoute le vecteur nul. L'ensemble des valeurs propres de u est appelé le *spectre* de u, et est noté $\mathrm{Sp}_K(u)$. Il peut être éventuellement vide.

Si $M \in \mathrm{M}_n(K)$ est une matrice carrée, on définit ses vecteurs propres et valeurs propres comme étant ceux de l'endomorphisme

$$\begin{aligned} K^n &\longrightarrow K^n \\ X &\longmapsto MX. \end{aligned}$$

Les endomorphismes possédant une matrice représentative diagonale dans une base bien choisie méritent un nom particulier.

2.3. Définition. Soit $u \in \mathscr{L}(E)$. On dit que u est *diagonalisable* s'il existe une base de E dans laquelle la matrice représentative de u soit diagonale. Cela revient à dire qu'il existe une base de E formée de vecteurs propres de u, ou encore que E se décompose en somme directe de droites stables par u.

Avant d'étudier l'existence ou non d'une base de vecteurs propres, il faut s'intéresser de plus près aux éléments propres. Le lemme suivant permet de trouver les valeurs propres d'un endomorphisme.

2.4. Lemme. *Soit $u \in \mathscr{L}(E)$. Alors, $\lambda \in K$ est une valeur propre de u si, et seulement si, $\chi_u(\lambda) = 0$.*

En particulier, $\mathrm{Sp}_K(u)$ est un ensemble fini.

Démonstration. Soit $\lambda \in K$. Alors, λ est une valeur propre de u si, et seulement si, $\mathrm{Ker}(\lambda \mathrm{Id}_E - u) \neq \{0\}$, ce qui revient à dire que $\lambda \mathrm{Id}_E - u$ est non inversible, ou encore que $\det(\lambda \mathrm{Id}_E - u) = 0$. Cela est équivalent à dire que $\chi_u(\lambda) = 0$.

Le dernier point provient du fait qu'un polynôme non nul de $K[X]$ possède un nombre fini de racines dans K. □

2.5. Définition. Soit $u \in \mathscr{L}(E)$, et soit $\lambda \in \mathrm{Sp}_K(u)$. La *multiplicité* de λ est sa multiplicité en tant que racine de χ_u. Autrement dit, c'est le plus grand entier $m \geqslant 1$ tel que $(X-\lambda)^m$ divise χ_u.

2.6. Proposition. *Soit $u \in \mathscr{L}(E)$. Pour tout $\lambda \in \mathrm{Sp}_K(u)$, on a*

$$1 \leqslant \dim_K(E_\lambda) \leqslant m_\lambda,$$

où m_λ est la multiplicité de λ.

Démonstration. Le fait que $1 \leqslant \dim_K(E_\lambda)$ provient des définitions (il existe au moins un vecteur propre non nul associé à λ). Soit $d = \dim_K(E_\lambda)$, et soit $(e_1, \ldots, e_d)$ une base de E_λ, que l'on complète en une base $\mathbf{e}$ de E. Alors, on a

$$\mathrm{Mat}(u\,;\mathbf{e}) = \begin{pmatrix} \lambda \mathrm{I}_d & B \\ 0 & C \end{pmatrix},$$

et donc $\chi_u = (X-\lambda)^d \chi_C$. Ainsi, λ est une racine d'ordre au moins d de χ_u, et donc $m_\lambda \geqslant d$. □

On cherche maintenant des conditions d'existence d'une base dans laquelle la matrice représentative de u soit diagonale. On commence par une conséquence directe du lemme de décomposition des noyaux.

2.7. Lemme. *Soit $u \in \mathscr{L}(E)$. Alors, les sous-espaces propres de u sont en somme directe.* □

Donnons un premier théorème de diagonalisation.

2.8. Théorème. *Les propriétés suivantes sont équivalentes :*

(1) *l'endomorphisme $u \in \mathscr{L}(E)$ est diagonalisable ;*

(2) *on a $E = \bigoplus_{\lambda \in \mathrm{Sp}_K(u)} E_\lambda$;*

(3) *le polynôme χ_u est scindé, et pour tout $\lambda \in \mathrm{Sp}_K(u)$, on a*

$$\dim_K(E_\lambda) = m_\lambda,$$

où m_λ est la multiplicité de λ.

Démonstration. On commence par montrer (3) $\Longrightarrow$ (2). Puisque χ_u est scindé, et que les racines de χ_u sont exactement les valeurs propres de u par le lemme 2.4, on a

$$\chi_u = \prod_{\lambda \in \mathrm{Sp}_K(u)} (X-\lambda)^{m_\lambda},$$

où m_λ est alors la multiplicité de λ par définition.

En particulier, en comparant les degrés, on obtient $\sum_{\lambda\in \mathrm{Sp}_K(u)} m_\lambda = n$. Mais alors, on a

$$\dim_K\Big(\bigoplus_{\lambda\in \mathrm{Sp}_K(u)} E_\lambda\Big) = \sum_{\lambda\in \mathrm{Sp}_K(u)} \dim_K(E_\lambda) = \sum_{\lambda\in \mathrm{Sp}_K(u)} m_\lambda = n = \dim_K(E).$$

On en déduit donc que $E = \bigoplus_{\lambda\in \mathrm{Sp}_K(u)} E_\lambda$.

L'implication (2) $\Longrightarrow$ (1) est claire, puisque en recollant une base de chaque sous-espace E_λ, on obtient une base de E formée de vecteurs propres de u.

Montrons enfin (1) $\Longrightarrow$ (3). Soit $\mathbf{e}$ une base de E formée de vecteurs propres de u. Quitte à renuméroter les vecteurs de $\mathbf{e}$, on peut toujours supposer que l'on a

$$M = \mathrm{Mat}(u\,;\mathbf{e}) = \begin{pmatrix} \lambda_1 \mathrm{I}_{m_1} & & \\ & \ddots & \\ & & \lambda_r \mathrm{I}_{m_r} \end{pmatrix},$$

où les λ_i sont distincts. Alors, on a

$$\chi_u = \prod_{i=1}^{r} (X-\lambda_i)^{m_i},$$

donc χ_u est scindé, et les valeurs propres distinctes de u sont $\lambda_1,\dots,\lambda_r$, de multiplicités respectives $m_1,\dots,m_r$. Par définition, les m_i vecteurs de base correspondant à la valeur propre λ_i sont des éléments de E_{λ_i}. On a donc $\dim_K(E_{\lambda_i}) \geqslant m_i$. Mais on sait aussi que $\dim_K(E_{\lambda_i}) \leqslant m_i$ par la proposition 2.6. Cela achève la démonstration. □

Ce théorème a un intérêt plutôt théorique. En pratique, il n'est pas très facile d'utilisation, puisqu'il faut calculer la dimension de tous les sous-espaces propres. Il a pourtant une conséquence intéressante.

2.9. Corollaire. *Soit $u \in \mathscr{L}(E)$. Si u possède $n = \dim_K(E)$ valeurs propres distinctes, alors u est diagonalisable.*

Démonstration. Il suffit d'appliquer la proposition 2.6 et le point (3) du théorème précédent. □

Nous allons maintenant donner un autre critère de diagonalisabilité. On commence par un lemme très simple, mais fort utile.

2.10. Lemme. *Soit $u \in \mathscr{L}(E)$, et soit $x \in E$ un vecteur propre de u associé à la valeur propre λ. Alors, pour tout $P \in K[X]$, on a*

$$P(u)(x) = P(\lambda)x.$$

En particulier, si P annule u, alors λ est une racine de P dans K.

Démonstration. On montre aisément par récurrence $u^m(x) = \lambda^m x$ pour tout $m \geqslant 0$. Si maintenant $P = a_d X^d + \cdots + a_1 X + a_0$, alors

$$P(u)(x) = (a_d u^d + \cdots + a_1 u + a_0 \mathrm{Id}_E)(x) = a_d \lambda^d x + \cdots + a_1 \lambda + a_0 x = P(\lambda)x.$$

Enfin, si $P(u) = 0$, on obtient $P(\lambda)x = 0$, d'où $P(\lambda) = 0$, car un vecteur propre est non nul. □

On peut maintenant énoncer et démontrer le théorème suivant.

2.11. Théorème. *Soit $u \in \mathscr{L}(E)$. Alors, les propriétés suivantes sont équivalentes :*

(1) *l'endomorphisme u est diagonalisable ;*

(2) *le polynôme $\prod_{\lambda \in \mathrm{Sp}_K(u)} (X - \lambda)$ annule u ;*

(3) *il existe un polynôme P annulant u scindé sans racines multiples. Dans ce cas, $\mathrm{Sp}_K(u)$ est contenu dans l'ensemble des racines de P ;*

(4) *le polynôme minimal μ_u est scindé sans racines multiples ;*

(5) $\mu_u = \prod_{\lambda \in \mathrm{Sp}_K(u)} (X - \lambda)$.

Démonstration

(1) $\Longrightarrow$ (2). Par le théorème 2.8, on a

$$E = \bigoplus_{\lambda \in \mathrm{Sp}_K(u)} E_\lambda = \bigoplus_{\lambda \in \mathrm{Sp}_K(u)} \mathrm{Ker}(u - \lambda \mathrm{Id}_E).$$

Le lemme de décomposition des noyaux donne alors $E = \mathrm{Ker}\big(P(u)\big)$, avec

$$P = \prod_{\lambda \in \mathrm{Sp}_K(u)} (X - \lambda),$$

qui est scindé sans racines multiples. Mais alors, on a $P(u) = 0$, et donc P annule u.

(2) $\Longrightarrow$ (3). La première partie est évidente, et la deuxième découle du lemme 2.10.

(3) $\Longrightarrow$ (4). Soit P un polynôme annulateur de u scindé sans racines multiples. Alors, $\mu_u \mid P$, et donc μ_u est a fortiori scindé sans racines multiples.

(4) $\Longrightarrow$ (5). On sait que μ_u et χ_u ont les mêmes racines dans K par la proposition 1.19. Les racines de μ_u sont donc exactement les valeurs propres de u par le lemme 2.4. Comme ces racines sont simples par hypothèse, on obtient le résultat.

(5) $\Longrightarrow$ (1). Supposons que $\mu_u = \prod_{\lambda \in \mathrm{Sp}_K(u)} (X - \lambda)$.

Comme μ_u annule u, le lemme de décomposition des noyaux donne

$$E = \bigoplus_{\lambda \in \mathrm{Sp}_K(u)} \mathrm{Ker}(u - \lambda \mathrm{Id}_E) = \bigoplus_{\lambda \in \mathrm{Sp}_K(u)} E_\lambda.$$

On en déduit alors que u est diagonalisable par le théorème 2.8. □

2.12. Remarques

(1) Le critère le plus utile pour démontrer la diagonalisabilité de u est certainement le point (3), car il ne présuppose pas de savoir calculer les valeurs propres. En revanche, il ne donne pas un moyen de trouver une base de vecteurs propres. Si l'on veut une telle base, il est indispensable de déterminer une base de chaque sous-espace propre, en résolvant des systèmes linéaires.

(2) Pour démontrer qu'un endomorphisme $u \in \mathscr{L}(E)$ n'est **pas** diagonalisable, on dispose de quelques méthodes. Soit on met en défaut le point (3) du théorème 2.8 en calculant la dimension des sous-espaces propres et en comparant avec les multiplicités des valeurs propres, soit on montre que μ_u n'est pas scindé à racines simples, soit on montre que le polynôme $\prod_{\lambda \in \mathrm{Sp}_K(u)} (X - \lambda)$ n'annule pas u.

On finit ces considérations sur la diagonalisabilité par un théorème de diagonalisation simultanée. On commence par un lemme, déjà intéressant par lui-même.

2.13. Lemme. *Soit $u \in \mathscr{L}(E)$, et soit F un sous-espace vectoriel stable par u. Si u est diagonalisable, alors $u_F \in \mathscr{L}(F)$ est diagonalisable.*

Démonstration. Si u est diagonalisable, alors μ_u est scindé sans racines multiples par le théorème précédent. Mais alors on a

$$\mu_u(u_F) = \mu_u(u)_F = 0,$$

et μ_u est donc un polynôme scindé sans racines multiples annulant u_F. Ainsi, u_F est diagonalisable par le théorème 2.11 (3). □

2.14. Théorème. *Soient $u, u' \in \mathscr{L}(E)$ tels que $u \circ u' = u' \circ u$. S'ils sont diagonalisables, alors ils sont diagonalisables dans une même base.*

Démonstration. Soit $\lambda \in K$ une valeur propre de u. Alors, E_λ est stable par u'. En effet, si $x \in E_\lambda$, on a

$$u\big(u'(x)\big) = u'\big(u(x)\big) = u'(\lambda x) = \lambda u'(x).$$

Par le lemme précédent, u'_{E_λ} est diagonalisable. Ainsi E_λ admet une base formée de vecteurs propres de u'_{E_λ}, qui sont a fortiori des vecteurs propres de u'. Mais les vecteurs de cette base sont également des vecteurs propres de u, puisque ce sont des éléments de E_λ. Comme u est diagonalisable, E est somme directe des E_λ. En recollant les bases précédentes, on obtient alors une base de E qui est constituée de vecteurs qui sont à la fois des vecteurs propres de u et de u'. □

3. Trigonalisation, décomposition de Dunford

Au vu des critères donnés dans la section précédente, on voit que, étant donné un endomorphisme $u \in \mathscr{L}(E)$, l'existence d'une décomposition

$$E = E_1 \oplus \cdots \oplus E_r$$

en somme directe de droites stables par u n'est pas assurée. Il faut donc être un peu plus raisonnable, et imposer des conditions moins fortes.

Par exemple, on pourrait se demander s'il existe des bases $\mathbf{e}_1, \ldots, \mathbf{e}_r$ des sous-espaces $E_1, \ldots, E_r$ respectivement telles que $\mathrm{Mat}(u_{E_i}; \mathbf{e}_i)$ soit triangulaire pour tout $i \in [\![1, n]\!]$.

En particulier, en recollant ces bases, on obtiendrait une base $\mathbf{e}$ de E telle que $\mathrm{Mat}(u\,; \mathbf{e})$ soit triangulaire.

Nous allons donc maintenant étudier l'existence de telles bases.

3.1. Définition. Soit $u \in \mathscr{L}(E)$. On dit que u est *trigonalisable* s'il existe une base $\mathbf{e}$ telle que $\mathrm{Mat}(u\,; \mathbf{e})$ soit triangulaire supérieure.

En guise d'exercice, le lecteur pourra vérifier que $u \in \mathscr{L}(E)$ est trigonalisable si, et seulement si, il existe des sous-espaces

$$\{0\} = E_0 \subset E_1 \subset E_2 \subset \cdots \subset E_n = E$$

stables par u tels que $\dim_K(E_i) = i$, pour $i \in [\![0, n]\!]$.

3.2. Remarque. Si $M = (a_{ij})$ est triangulaire supérieure, alors il est facile de voir que l'on a

$$\chi_M = \prod_{i=1}^{n} (X - a_{ii}),$$

et les termes diagonaux a_{ii} sont alors les valeurs propres de M. En particulier, si u est trigonalisable, χ_u est scindé sur K.

3.3. Théorème. *Soit $u \in \mathscr{L}(E)$. Alors, u est trigonalisable si, et seulement si, son polynôme caractéristique est scindé sur K.*

Dans ce cas, si $\lambda_1, \ldots, \lambda_r \in K$ sont les valeurs propres distinctes de u, il existe une base $\mathbf{e}$ de E telle que

$$\mathrm{Mat}(u\,;\mathbf{e}) = \begin{pmatrix} \lambda_1 & * & \cdots & \cdots & \cdots & \cdots & * \\ & \ddots & * & & & & * \\ & & \lambda_1 & * & & & * \\ & & & \ddots & \ddots & & \vdots \\ & & & & \lambda_r & \ddots & \vdots \\ & & & & & \ddots & * \\ & & & & & & \lambda_r \end{pmatrix}.$$

Démonstration. On a déjà vu que la condition était nécessaire. Pour montrer qu'elle est suffisante, nous allons démontrer l'existence dune base $\mathbf{e}$ possédant les propriétés voulues par récurrence sur $n = \dim_K(E)$. Pour tout $n \geqslant 1$, soit (H_n) la propriété :

(H_n) Pour tout K-espace vectoriel E de dimension n, et tout $u \in \mathscr{L}(E)$, il existe une base $\mathbf{e}$ de E vérifiant les propriétés voulues.

La propriété (H_1) est clairement vérifiée. Supposons que (H_n) soit vraie au rang $n \geqslant 1$, et montrons (H_{n+1}).

Soit E un K-espace vectoriel de dimension $n+1$ et soit $u \in \mathscr{L}(E)$ tel que le polynôme χ_u soit scindé sur K. Écrivons

$$\chi_u = \prod_{i=1}^{r} (X - \lambda_i)^{m_i},$$

où $\lambda_1, \ldots, \lambda_r \in K$ sont les valeurs propres distinctes de u. Soit $x \in E \setminus \{0\}$ un vecteur propre de u associé à la valeur propre λ_1. Complétons ce vecteur en une base $\mathbf{e} = (x, e'_1, \ldots, e'_n)$ de E. Puisque l'on a l'égalité $u(x) = \lambda_1 x$, la matrice $M = \mathrm{Mat}(u\,;\mathbf{e})$ est de la forme

$$M = \begin{pmatrix} \lambda_1 & * \\ 0 & M' \end{pmatrix}.$$

De plus, on a $\chi_M = (X - \lambda_1)\chi_{M'}$, et par conséquent

$$\chi_{M'} = (X - \lambda_1)^{m_1 - 1} \prod_{i=2}^{r} (X - \lambda_i)^{m_i}.$$

Ainsi, $\chi_{M'}$ est scindé.

Soit $E' = \mathrm{Vect}(e'_1, \ldots, e'_n)$ et soit $u' \in \mathscr{L}(E')$ l'unique endomorphisme de E' dont la matrice représentative dans la base $\mathbf{e}' = (e'_1, \ldots, e'_n)$ est M'.

Alors, E' est de dimension n, et $\chi_{u'}$ est scindé. Notons que si $m_1 > 1$, les valeurs propres de M' sont $\lambda_1, \ldots, \lambda_r$, et si $m_1 = 1$, les valeurs propres de M' sont $\lambda_2, \ldots, \lambda_r$.

Par hypothèse de récurrence, il existe une base $\mathbf{f} = (f_2, \ldots, f_{n+1})$ de E' telle que

$$T = \mathrm{Mat}(u'; \mathbf{f}) = \begin{pmatrix} \lambda_1 & * & \cdots & \cdots & \cdots & \cdots & * \\ & \ddots & * & & & & * \\ & & \lambda_1 & * & & & * \\ & & & \ddots & \ddots & & \vdots \\ & & & & \lambda_r & \ddots & \vdots \\ & & & & & \ddots & * \\ & & & & & & \lambda_r \end{pmatrix}$$

si $m_1 > 1$, et

$$T = \mathrm{Mat}(u'; \mathbf{f}) = \begin{pmatrix} \lambda_2 & * & \cdots & \cdots & \cdots & \cdots & * \\ & \ddots & * & & & & * \\ & & \lambda_2 & * & & & * \\ & & & \ddots & \ddots & & \vdots \\ & & & & \lambda_r & \ddots & \vdots \\ & & & & & \ddots & * \\ & & & & & & \lambda_r \end{pmatrix}$$

si $m_1 = 1$.

Soit P' la matrice de passage de $\mathbf{f}$ à $\mathbf{e}'$. Alors, on a l'égalité $T = P'^{-1} M' P'$. Soit $\mathbf{t} = (x, f_2, \ldots, f_{n+1})$. La matrice de passage de $\mathbf{t}$ à $\mathbf{e}$ est

$$Q = \begin{pmatrix} 1 & 0 \\ 0 & P' \end{pmatrix},$$

et l'on a

$$\mathrm{Mat}(u; \mathbf{t}) = Q^{-1} M Q = \begin{pmatrix} \lambda_1 & * \\ 0 & T \end{pmatrix},$$

qui est de la forme voulue dans les deux cas. Ainsi, (H_{n+1}) est vraie. Cela achève la récurrence, ainsi que la démonstration. □

Il découle par exemple de ce théorème que tout endomorphisme d'un espace vectoriel complexe de dimension finie est trigonalisable. Autrement dit, pour toute matrice $A \in \mathrm{M}_n(\mathbb{C})$, il existe une matrice $P \in \mathrm{GL}_n(\mathbb{C})$ et une matrice triangulaire $T \in \mathrm{M}_n(\mathbb{C})$ telles que $A = PTP^{-1}$. Malheureusement, le calcul des puissances d'une matrice triangulaire, même s'il est a priori plus simple que dans le cas d'une matrice quelconque, n'est pas chose si aisée.

Il y a néanmoins deux cas particuliers où c'est particulièrement facile : le cas où T est diagonale/diagonalisable (comme on l'a déjà vu dans la section précédente), et le cas où T est à diagonale nulle. Dans ce dernier cas, le théorème de Cayley-Hamilton montre que $T^n = 0$, et donc $T^m = 0$ pour tout $m \geqslant n$.

Si T est une matrice $n \times n$ triangulaire quelconque, alors T s'écrit sous la forme $T = D + N$, où D est diagonale, et où N est triangulaire à diagonale nulle. Dans le cas favorable où D et N commutent, on peut utiliser la formule du binôme de Newton, ainsi que le fait que $N^k = 0$ pour tout entier $k \geqslant n$ pour obtenir

$$T^m = \sum_{k=0}^{n-1} \binom{m}{k} D^{m-k} N^k, \qquad \text{pour tout } m \geqslant 0,$$

ce qui est facile à calculer.

Le but de ce qui va suivre est démontrer que l'on peut toujours se ramener à un tel cas, après changement de base approprié.

On commence par une définition.

3.4. Définition. Un endomorphisme $\nu \in \mathscr{L}(E)$ est dit *nilpotent* s'il existe un entier $k \geqslant 1$ tel que $\nu^k = 0$. *L'indice de nilpotence* de ν, noté n_ν, est alors le plus petit entier $k \geqslant 1$ satisfaisant $\nu^k = 0$.

Une matrice $N \in \mathrm{M}_n(K)$ est dite *nilpotente* s'il existe un entier $k \geqslant 1$ tel que $N^k = 0$. On définit alors comme précédemment son indice de nilpotence. De plus, un endomorphisme $\nu \in \mathscr{L}(E)$ est nilpotent si, et seulement si, sa matrice représentative dans une base arbitraire est nilpotente.

Nous allons maintenant étudier quelques propriétés élémentaires de ces endomorphismes, que nous regroupons dans un lemme.

3.5. Lemme. *Soit $\nu \in \mathscr{L}(E)$ un endomorphisme. Alors, les propriétés suivantes sont équivalentes :*

(1) *l'endomorphisme ν est nilpotent ;*

(2) *on a $\mu_\nu = X^r$ pour un certain $r \geqslant 1$;*
Dans ce cas, $n_\nu = r$. En particulier, $n_\nu \leqslant \dim_K(E)$.

(3) *le polynôme χ_ν est scindé et $\mathrm{Sp}_K(\nu) = \{0\}$;*

(4) *on a $\chi_\nu = X^n$.*

Enfin, on a les assertions suivantes :

(i) *le seul endomorphisme à la fois diagonalisable et nilpotent est l'endomorphisme nul ;*

(ii) *la somme de deux endomorphismes nilpotents qui commutent est encore un endomorphisme nilpotent.*

Démonstration

(1) $\Longrightarrow$ (2). Si ν est nilpotent, on a $\nu^k = 0$ pour un certain $k \geqslant 1$. Alors, le polynôme X^k annule ν, et μ_ν est donc une puissance non triviale de X, puisque $\mu_\nu \mid X^k$. On a donc $\mu_u = X^r$, avec $r \geqslant 1$. Comme $u^{n_\nu} = 0$, X^{n_ν} annule ν, et par suite $\mu_\nu \mid X^{n_\nu}$. Ainsi, $r \leqslant n_\nu$. Mais, par définition de n_ν, on a $\nu^i \neq 0$ pour $i \in [\![1, n_\nu - 1]\!]$. Puisque $\mu_\nu(\nu) = 0 = \nu^r$, on a donc $r \geqslant n_\nu$, d'où $n_\nu = r$. Enfin, comme $\mu_\nu \mid \chi_\nu$, $\deg(\mu_\nu) = r \leqslant \deg(\chi_\nu) = \dim_K(E)$.

(2) $\Longrightarrow$ (3). Comme μ_ν et χ_ν ont les mêmes facteurs irréductibles par la proposition 1.19, on en déduit que $\chi_\nu = X^n$. Ainsi, χ_ν est scindé et $\mathrm{Sp}_K(\nu) = \{0\}$.

(3) $\Longrightarrow$ (4). Si χ_ν est scindé et $\mathrm{Sp}_K(\nu) = \{0\}$, alors $\chi_\nu = X^n$, puisque les racines de χ_ν sont les valeurs propres de ν.

(4) $\Longrightarrow$ (1). Supposons que $\chi_\nu = X^n$. Comme χ_ν annule ν, on en déduit que $\nu^n = 0$, ce qui montre que ν est nilpotent.

Cela démontre l'équivalence des quatre propriétés.

Montrons (i). Par (3), un endomorphisme nilpotent ν n'a que 0 comme valeur propre. Mais si ν est diagonalisable, il existe alors une base dans laquelle sa matrice représentative est nulle. Ainsi, $\nu = 0$. Inversement, l'endomorphisme nul est bien nilpotent et diagonalisable.

Il reste à montrer (ii). Soient ν et ν' deux endomorphismes nilpotents de E qui commutent. Soient k, $\ell \geqslant 1$ tels que $\nu^k = \nu'^\ell = 0$. Puisque ν et ν' commutent, on a alors

$$(\nu + \nu')^{k+\ell} = \sum_{i=0}^{k+\ell} \binom{k+\ell}{i} \nu^i \circ (\nu')^{k+\ell-i}.$$

Si $i \geqslant k$, alors $\nu^i = 0$. Si $i \leqslant k$, alors $k + \ell - i \geqslant \ell$ et donc $(\nu')^{k+\ell-i} = 0$. On a donc

$$(\nu + \nu')^{k+\ell} = 0,$$

et $\nu + \nu'$ est nilpotent. □

Revenons à notre problème. Soit $u \in \mathscr{L}(E)$ de polynôme caractéristique scindé $\chi_u = \prod_{i=1}^{r} (X - \lambda_i)^{m_i}$. Soit $\mathbf{t}$ une base de E telle que $T = \mathrm{Mat}(u\,;\mathbf{t})$ soit triangulaire supérieure. On peut même supposer que les coefficients diagonaux de T sont $\lambda_1, \ldots, \lambda_1, \ldots, \lambda_r, \ldots, \lambda_r$, ce que nous ferons désormais. Ainsi, on peut écrire T sous la forme d'une matrice triangulaire supérieure par blocs $T = (T_{ij})_{i,j}$, avec $T_{ij} \in \mathrm{M}_{m_i \times m_j}(K)$ et où tous les éléments diagonaux de T_{ii} sont égaux à λ_i.

Posons $N_i = T_{ii} - \lambda_i \mathrm{I}_{m_i}$. Remarquons que la matrice N_i est triangulaire supérieure à diagonale nulle.

Soient

$$D = \begin{pmatrix} \lambda_1 \mathrm{I}_{m_1} & & & \\ & \ddots & & \\ & & \ddots & \\ & & & \lambda_r \mathrm{I}_{m_r} \end{pmatrix} \quad \text{et} \quad N = \begin{pmatrix} N_1 & & & \\ & \ddots & T_{ij} & \\ & & \ddots & \\ & & & N_r \end{pmatrix}$$

On a donc $T = D + N$. Supposons que D et N commutent. Alors, on a facilement $\lambda_i T_{ij} = \lambda_j T_{ij}$, pour tous $i, j \in [\![1, r]\!]$ vérifiant $i > j$. Puisque l'on a $\lambda_i \neq \lambda_j$, on obtient $T_{ij} = 0$. Ainsi, T est diagonale par blocs et chaque bloc diagonal est de la forme

$$T_{ii} = \lambda_i \mathrm{I}_{m_i} + N_i,$$

où $N_i \in \mathrm{M}_{m_i}(K)$ est triangulaire supérieure à diagonale nulle. En particulier, $N_i^{m_i} = 0$ et T_{ii} possède une unique valeur propre, à savoir λ_i.

Pour $k \in [\![1, r]\!]$ fixé, la matrice $T - \lambda_k \mathrm{I}_n$ est alors diagonale par blocs, où le i-ième bloc est de la forme

$$\begin{pmatrix} \lambda_i - \lambda_k & * & * \\ & \ddots & * \\ & & \lambda_i - \lambda_k \end{pmatrix}.$$

Un calcul par blocs montre alors que $\mathrm{Ker}\big((u - \lambda_k \mathrm{Id}_E)^{m_k}\big)$ est engendré par les éléments de $\mathbf{t}$ correspondants aux indices de la « k-ième tranche verticale » de T.

Autrement dit, $\mathbf{t}$ est nécessairement obtenue en recollant des bases appropriées de $\mathrm{Ker}\big((u - \lambda_i \mathrm{Id}_E)^{m_i}\big)$, $i \in [\![1, r]\!]$.

Tout cela nous conduit naturellement à étudier ces sous-espaces.

3.6. Définition. Soit $u \in \mathscr{L}(E)$, et soit $\lambda \in \mathrm{Sp}_K(u)$, de multiplicité m_λ. *Le sous-espace caractéristique* de u associé à λ est le sous-espace

$$E_\lambda^{m_\lambda} = \mathrm{Ker}\big((u - \lambda \mathrm{Id}_E)^{m_\lambda}\big).$$

Remarquons que l'on ne suppose pas ici que χ_u est scindé.

Le résultat suivant résume les propriétés importantes du sous-espace caractéristique.

3.7. Proposition. *Soit $u \in \mathscr{L}(E)$, et soit $\lambda \in \mathrm{Sp}_K(u)$. Soit m_λ la multiplicité de λ et soit r_λ la multiplicité de λ comme racine de μ_u.*

Alors, le sous-espace caractéristique $E_\lambda^{m_\lambda}$ est stable par u, de dimension m_λ. De plus, $u_\lambda = (u - \lambda \mathrm{Id}_E)_{E_\lambda^{m_\lambda}} \in \mathscr{L}(E_\lambda^{m_\lambda})$ est nilpotente, d'indice de nilpotence r_λ.

Démonstration. Écrivons $\chi_u = (X-\lambda)^{m_\lambda} R$, où $R \in K[X]$ n'est pas divisible par $X - \lambda$. En particulier, les polynômes $X - \lambda$ et R sont premiers entre eux. L'exemple 1.3 nous montre alors que $E_\lambda^{m_\lambda}$ est stable par u et le lemme 1.23 donne $\dim_K(E_\lambda^{m_\lambda}) = m_\lambda$.

Si l'on écrit cette fois $\mu_u = (X-\lambda)^{r_\lambda} S$, où $S \in K[X]$ n'est pas divisible par $X - \lambda$, le corollaire 1.21 donne l'égalité

$$\mu_{u_{E_\lambda^{m_\lambda}}} = (X-\lambda)^{r_\lambda}.$$

Notons que pour tout $P \in K[X]$, on a

$$P(u_\lambda) = 0 \iff P\big(u_{E_\lambda^{m_\lambda}} - \lambda \mathrm{Id}_{E_\lambda^{m_\lambda}}\big) = 0 \iff \mu_{u_{E_\lambda^{m_\lambda}}} \mid P(X-\lambda).$$

Ainsi, pour tout $P \in K[X]$, on a

$$P(u_\lambda) = 0 \iff \mu_{u_{E_\lambda^{m_\lambda}}}(X+\lambda) \mid P \iff X^{r_\lambda} \mid P.$$

On en déduit que $\mu_{u_\lambda} = X^{r_\lambda}$. Par le lemme 3.5, u_λ est nilpotent, d'indice de nilpotence r_λ. □

Supposons maintenant que χ_u soit scindé, et écrivons

$$\chi_u = \prod_{\lambda \in \mathrm{Sp}_K(u)} (X-\lambda)^{m_\lambda}.$$

Par le lemme de décomposition des noyaux, on a

$$E = \bigoplus_{\lambda \in \mathrm{Sp}_K(u)} E_\lambda^{m_\lambda}.$$

3.8. Définition. Soit $u \in \mathscr{L}(E)$, de polynôme caractéristique scindé sur K. On appelle *projecteur spectral* associé à $\lambda \in \mathrm{Sp}_K(u)$ l'unique endomorphisme $p_\lambda \in \mathscr{L}(E)$ vérifiant les deux conditions suivantes :

(1) $p_\lambda(x) = x$, pour tout $x \in E_\lambda^{m_\lambda}$;

(2) $p_\lambda(x) = 0$, pour tout $x \in E_\mu^{m_\mu}$ et tout $\mu \neq \lambda$.

Autrement dit, p_λ est la projection sur $E_\lambda^{m_\lambda}$ parallèlement à $\bigoplus_{\mu \neq \lambda} E_\mu^{m_\mu}$.

On va maintenant regarder d'un peu plus près ces projecteurs spectraux, qui interviennent (souvent de manière un peu cachée) dans la démonstration de l'existence d'une décomposition de Dunford.

3.9. Lemme. *Soit $u \in \mathscr{L}(E)$ tel que χ_u soit scindé. Alors, il existe des polynômes $U_\lambda \in K[X]$ tels que*

$$\frac{1}{\chi_u} = \sum_{\lambda \in \mathrm{Sp}_K(u)} \frac{U_\lambda}{(X-\lambda)^{m_\lambda}}.$$

De plus, le projecteur spectral associé à λ est

$$p_\lambda = (Q_\lambda U_\lambda)(u),$$

avec

$$Q_\lambda = \frac{\chi_u}{(X-\lambda)^{m_\lambda}} = \prod_{\mu \neq \lambda} (X-\mu)^{m_\mu}, \quad \textit{pour tout } \lambda \in \mathrm{Sp}_K(u).$$

Démonstration. En réduisant au même dénominateur les fractions correspondant au pôle λ dans la décomposition en éléments simples de $\frac{1}{\chi_u}$, on obtient une égalité de la forme

$$\frac{1}{\chi_u} = \sum_{\lambda \in \mathrm{Sp}_K(u)} \frac{U_\lambda}{(X-\lambda)^{m_\lambda}},$$

où $U_\lambda \in K[X]$.

En réduisant maintenant toutes ces fractions au même dénominateur, on obtient $\frac{1}{\chi_u} = \sum_{\lambda \in \mathrm{Sp}_K(u)} \frac{Q_\lambda U_\lambda}{\chi_u}$, d'où

$$\sum_{\lambda \in \mathrm{Sp}_K(u)} Q_\lambda U_\lambda = 1.$$

Si $\mu \neq \lambda$, alors $(X-\mu)^{m_\mu} \mid Q_\lambda$, et par conséquent $(Q_\lambda U_\lambda)(u)(x) = 0$ pour tout $x \in E_\mu^{m_\mu} = \mathrm{Ker}\big((u-\mu \mathrm{Id}_E)^{m_\mu}\big)$. Mais $\sum_{\mu \in \mathrm{Sp}_K(u)} Q_\mu U_\mu = 1$, et donc pour tout $x \in E_\lambda^{m_\lambda}$, on a

$$\Big(\sum_{\mu \in \mathrm{Sp}_K(u)} Q_\mu U_\mu\Big)(u)(x) = (Q_\lambda U_\lambda)(u)(x) = \mathrm{Id}_E(x) = x.$$

Ainsi, $p_\lambda = (Q_\lambda U_\lambda)(u)$, et cela achève la démonstration. □

On va maintenant démontrer le résultat suivant.

3.10. Théorème (Dunford). *Soit $u \in \mathscr{L}(E)$. On suppose que χ_u est scindé. Alors, il existe des endomorphismes $\delta, \nu \in \mathscr{L}(E)$ tels que*

(1) *δ et ν commutent ;*

(2) *δ est diagonalisable et ν est nilpotent ;*

(3) *$u = \delta + \nu$.*

De plus, une telle décomposition est unique. Plus précisément, on a

$$\delta = \sum_{\lambda \in \mathrm{Sp}_K(u)} \lambda p_\lambda,$$

où p_λ est le projecteur spectral associé à λ.

Enfin, δ et ν sont des polynômes en u.

Démonstration. Posons

$$\delta = \sum_{\lambda \in \mathrm{Sp}_K(u)} \lambda p_\lambda \quad \text{et} \quad \nu = u - \delta,$$

et montrons que cette décomposition convient.

D'après le lemme 3.9, les projecteurs spectraux sont des polynômes en u. Il en est donc de même de δ, donc de $\nu = u - \delta$. En particulier, δ et ν commutent par le lemme 1.2. De plus, puisque E^{m_λ} est stable par u, il est stable par tout polynôme en u, donc stable par δ et ν.

La définition de p_λ montre alors que l'on a

$$\delta_{E_\lambda^{m_\lambda}} = \lambda \mathrm{Id}_{E_\lambda^{m_\lambda}}, \quad \text{pour tout } \lambda \in \mathrm{Sp}_K(u).$$

En particulier, en recollant des bases de $E_\lambda^{m_\lambda}$, $\lambda \in \mathrm{Sp}_K(u)$, on obtient une base de E dans laquelle la matrice représentative de δ est diagonale. Ainsi, δ est diagonalisable.

Par ailleurs,

$$\nu_{E_\lambda^{m_\lambda}} = u_{E_\lambda^{m_\lambda}} - \lambda \mathrm{Id}_{E_\lambda^{m_\lambda}} = (u - \lambda \mathrm{Id}_E)_{E_\lambda^{m_\lambda}}, \quad \text{pour tout } \lambda \in \mathrm{Sp}_K(u).$$

Par la proposition 3.7, on en déduit que $\nu_{E_\lambda^{m_\lambda}} \in \mathscr{L}(E_\lambda^{m_\lambda})$ est nilpotent. De plus, son indice de nilpotence est inférieur ou égal à $\dim_K(E_\lambda^{m_\lambda})$, donc à n.

Ainsi,

$$(\nu^n)_{E_\lambda^{m_\lambda}} = (\nu_{E_\lambda^{m_\lambda}})^n = 0, \quad \text{pour tout } \lambda \in \mathrm{Sp}_K(u).$$

Comme $E = \bigoplus_{\lambda \in \mathrm{Sp}_K(u)} E_\lambda^{m_\lambda}$, on en déduit que $\nu^n = 0$, et donc ν est nilpotent.

Supposons maintenant que l'on ait une autre décomposition

$$u = \delta' + \nu'.$$

On a donc $\delta - \delta' = \nu' - \nu$. Puisque les endomorphismes δ' et ν' commutent, δ' commute avec $\delta' + \nu' = u$. Ainsi, δ' commute avec tout polynôme en u. En particulier, δ et δ' commutent. De même, ν et ν' commutent. Comme δ et δ' sont des endomorphismes diagonalisables, ils le sont dans une même base par le théorème 2.14. Il est alors facile de voir que $\delta - \delta'$ est aussi diagonalisable (la différence de deux matrices diagonales est diagonale !).

Maintenant, puisque ν' et $-\nu$ sont nilpotents, $\nu' - \nu$ est encore nilpotent par le lemme 3.5. Ainsi,

$$\delta - \delta' = \nu' - \nu$$

est à la fois diagonalisable et nilpotent, et donc nul, d'après ce même lemme. On a donc $\delta' = \delta$, et ainsi $\nu' = \nu$. Cela achève la démonstration. □

3.11. Définition. Soit $u \in \mathscr{L}(E)$ tel que χ_u soit scindé.

La décomposition $u = \delta + \nu$ donnée par le théorème précédent est appelée la *décomposition de Dunford* de u.

3.12. Remarques

(1) La démonstration fournit une méthode pour trouver de manière explicite la décomposition de Dunford pourvu que l'on sache trouver les valeurs propres, c'est-à-dire factoriser χ_u.

On écrit $\chi_u = \prod_{\lambda \in \mathrm{Sp}_K(u)} (X - \lambda)^{m_\lambda}$. À l'aide de la décomposition en éléments simples de $\frac{1}{\chi_u}$, on trouve des polynômes $U_\lambda \in K[X]$ tels que

$$\frac{1}{\chi_u} = \sum_{\lambda \in \mathrm{Sp}_K(u)} \frac{U_\lambda}{(X-\lambda)^{m_\lambda}}.$$

Soit $Q_\lambda = \prod_{\mu \neq \lambda} (X - \mu)^{m_\mu}$.

On a alors

$$\delta = \Big(\sum_{\lambda \in \mathrm{Sp}_K(u)} \lambda Q_\lambda U_\lambda \Big)(u) \quad \text{et} \quad \nu = u - \delta.$$

(2) La démonstration fonctionne parfaitement si l'on remplace χ_u par un polynôme annulateur P tel que $P \mid \chi_u$, i.e. un polynôme annulateur de la forme

$$P = \bigoplus_{\lambda \in \mathrm{Sp}_K(u)} (X - \lambda)^{s_\lambda}, \quad s_\lambda \geqslant 1.$$

On a alors une décomposition

$$E = \prod_{\lambda \in \mathrm{Sp}_K(u)} \mathrm{Ker}\big((u - \lambda \mathrm{Id}_E)^{s_\lambda}\big).$$

On calcule ensuite les projecteurs associés comme expliqué dans le lemme 3.9, mais en utilisant P plutôt que χ_u. En pratique, on aura intérêt à faire les calculs avec μ_u plutôt que χ_u. De toute façon, avant tout calcul, il vaut mieux vérifier si u est diagonalisable ou pas et donc calculer μ_u. En effet, si u est diagonalisable, alors on a $\delta = u$ et $\nu = 0$.

On peut même se passer de l'hypothèse $P \mid \chi_u$, et prendre un polynôme annulateur P scindé sur K, sans avoir à déterminer quelles racines de P sont les valeurs propres. En effet, on sait déjà que toutes les valeurs propres de u sont racines de P par le lemme 2.10. D'autre part, si $\alpha \in K$ est une racine de P qui n'est pas une valeur propre de u, alors on a $\mathrm{Ker}\big((u-\alpha \mathrm{Id}_E)^{s_\alpha}\big) = \{0\}$, comme on peut le voir en écrivant la matrice représentative de $(u-\alpha \mathrm{Id}_E)^{s_\alpha}$ dans une base qui trigonalise l'endomorphisme u. Ainsi, les racines de P qui ne sont pas valeurs propres ne donnent que des sous-espaces nuls (et donc quelques calculs inutiles).

(3) La décomposition de Dunford a bien entendu son équivalent matriciel : pour toute matrice $A \in \mathrm{M}_n(K)$ de polynôme caractéristique scindé, il existe D, $N \in \mathrm{M}_n(K)$, avec D diagonalisable et N nilpotente, telles que

$$DN = ND \quad \text{et} \quad A = D + N.$$

De plus, D et N sont uniques, et sont des polynômes en A.

Nous ne résistons pas à faire une parenthèse pour donner une élégante méthode de diagonalisation utilisant les projecteurs spectraux.

3.13. Proposition. *Soit $u \in \mathscr{L}(E)$ un endomorphisme diagonalisable. Pour tout $\lambda \in \mathrm{Sp}_K(u)$, on pose*

$$L_\lambda = \prod_{\mu \neq \lambda} \frac{X-\mu}{\lambda-\mu}.$$

Alors, le projecteur p_λ sur le sous-espace propre E_λ associé à la décomposition $E = \bigoplus_{\lambda \in \mathrm{Sp}_K(u)} E_\lambda$ est donné par

$$p_\lambda = L_\lambda(u).$$

En particulier, si $\mathbf{e} = (e_1, \ldots, e_n)$ est une base de E fixée, une base du sous-espace propre E_λ est obtenue en extrayant m_λ vecteurs linéairement indépendants de la famille $\big(L_\lambda(u)(e_1), \ldots, L_\lambda(u)(e_n)\big)$.

Démonstration. Puisque u est diagonalisable, on a $\mu_u = \prod_{\lambda \in \mathrm{Sp}_K(u)} (X-\lambda)$ par le théorème 2.11. On applique la méthode décrite précédemment, avec le polynôme $P = \mu_u$. On pose

$$Q_\lambda = \prod_{\mu \neq \lambda} (X-\mu).$$

De plus, une décomposition en éléments simples de $\frac{1}{\mu_u}$ donne

$$\frac{1}{\prod_{\lambda \in \mathrm{Sp}_K(u)} (X-\lambda)} = \sum_{\lambda \in \mathrm{Sp}_K(u)} \frac{1}{\prod_{\mu \neq \lambda} (\lambda-\mu)} \times \frac{1}{X-\lambda}.$$

Par le lemme 3.9, on a alors $p_\lambda = L_\lambda(u)$. Le dernier point vient juste du fait que l'image de p_λ est E_λ, par définition même. □

3.14. Remarque. Matriciellement, le résultat se retraduit de la façon suivante : si $A = \mathrm{Mat}(u\,;\mathbf{e})$, une base de E_λ est obtenue en prenant m_λ colonnes linéairement indépendantes de la matrice

$$\Pi_\lambda = \mathrm{Mat}(p_\lambda\,;\mathbf{e}) = \prod_{\mu\neq\lambda} \frac{A-\mu \mathrm{I}_n}{\lambda-\mu}.$$

Cette méthode est d'une efficacité redoutable lorsque toutes les valeurs propres sont simples. Dans ce cas, pour trouver un vecteur propre associé à la valeur propre λ, il suffit de choisir un vecteur non nul parmi tous les vecteurs $L_\lambda(A)(e_i)$, $i \in [\![1,n]\!]$.

L'algorithme de calcul de la décomposition de Dunford proposé dans la remarque 3.12 (1) n'est pas entièrement satisfaisant, car il présuppose que l'on sache calculer les valeurs propres.

On va maintenant donner une méthode alternative pour calculer la décomposition de Dunford, et qui permet d'éviter ce calcul, au moins dans certains cas (voir la remarque 3.16).

3.15. Proposition. *Soit $u \in \mathscr{L}(E)$ tel que χ_u soit scindé, et soit*

$$P = \prod_{\lambda\in \mathrm{Sp}_K(u)} (X-\lambda).$$

La suite $(u_n)_{n\geqslant 0}$ d'éléments de $\mathscr{L}(E)$ vérifiant

$$u_0 = u, \quad u_{n+1} = u_n - P(u_n)\circ P'(u_n)^{-1} \in \mathscr{L}(E), \quad n \geqslant 0$$

est bien définie et stationnaire. Si $n \geqslant 1$ est un entier à partir duquel la suite stationne, et si l'on pose $\delta = u_n$, $\nu = u - u_n$, alors $u = \delta + \nu$ est la décomposition de Dunford de u.

Démonstration. Si $Q \in K[X]$, on note $Q(u)K[u]$ l'idéal de $K[u]$ engendré par $Q(u)$. Écrivons

$$\chi_u = \prod_{\lambda\in \mathrm{Sp}_K(u)} (X-\lambda)^{m_\lambda},$$

et soit $P = \prod_{\lambda\in \mathrm{Sp}_K(u)} (X-\lambda)$. On a donc $P \mid \chi_u$. Si $r = \max_{\lambda\in \mathrm{Sp}_K(u)} (m_\lambda)$, on a

$$P(u)^r = (P^r)(u) = 0,$$

car $\chi_u \mid P^r$, et donc $P(u)$ est nilpotent. Par ailleurs, P est à racines simples, donc P et P' sont sans racines communes, et donc P' et χ_u sont premiers

entre eux. On a alors une relation de Bézout

$$UP' + V\chi_u = 1, \quad U, V \in K[X],$$

et donc $U(u) \circ P'(u) = \mathrm{Id}_E$, car $\chi_u(u) = 0$, ce qui montre que $P'(u)$ est inversible, et que son inverse est un polynôme en u.

Cela étant dit, posons

$$u_0 = u, \quad u_{n+1} = u_n - P(u_n) \circ P'(u_n)^{-1} \in \mathscr{L}(E), \quad n \geqslant 0.$$

On va montrer par récurrence que cette définition a bien un sens, c'est-à-dire que $P'(u_n) \in \mathrm{GL}(E)$ pour tout $n \geqslant 0$. Plus précisément, soit (H_n) la liste de propriétés :

$$(H_n) \qquad u_n \in K[u], \quad P'(u_n) \in K[u]^\times, \quad P(u_n) \in P(u)^{2^n} K[u].$$

Supposons avoir montré que (H_n) est vraie pour tout entier $n \geqslant 0$, et achevons la démonstration du théorème. Remarquons que la suite stationne à partir du rang n si, et seulement si, $P(u_n) = 0$ $\big(\text{car } P'(u_n) \in \mathrm{GL}(E)\big)$. Puisque l'on a $P(u)^r = 0$, alors pour tout entier $m \geqslant 0$ tel que $2^m \geqslant r$, on a $P(u_m) = 0$ puisque $P(u_m) \in P(u)^{2^m} K[u]$, et la suite $(u_n)_{n \geqslant 0}$ est stationnaire. Choisissons un entier $n \geqslant 1$ à partir duquel la suite stationne, soit $\delta = u_n$ et soit $\nu = u - \delta$. Vérifions que δ et ν ont bien les propriétés requises.

Puisque $\delta = u_n = u_{n+1}$, on a $\delta \in K[u]$ et $P(\delta) = 0$. Comme P est scindé sans racines multiples, l'endomorphisme δ est alors diagonalisable par le théorème 2.11. De plus, $\nu \in K[u]$ puisque $\delta \in K[u]$. Il reste à voir que ν est nilpotent. Mais pour tout $j \geqslant 0$, on a

$$u_j - u_{j+1} \in P(u)^{2^j} K[u] \subset P(u)K[u].$$

En additionnant, on trouve que $\nu = u - \delta = u_0 - u_n \in P(u)K[u]$. L'endomorphisme $P(u)$ étant nilpotent, ν l'est également. En effet, si $R \in K[X]$, on a

$$\big(P(u) \circ R(u)\big)^r = P(u)^r \circ R(u)^r = 0,$$

la première égalité étant vraie car $P(u)$ et $R(u)$ commutent.

Il reste à montrer (H_n) pour tout $n \geqslant 0$. Nous allons procéder par récurrence sur n.

Puisque $u_0 = u$, que $P(u_0) = P(u)^1$, et que $P'(u)$ est inversible dans $K[u]$, la propriété (H_0) est vraie. Supposons que (H_n) soit vraie pour un certain entier $n \geqslant 0$, et montrons que (H_{n+1}) est vraie. Puisque $u_n \in K[u]$, il en est de même de $P(u_n)$ et de $P'(u_n)$. Par ailleurs, $P'(u_n)$ est inversible dans $K[u]$, donc $P'(u_n)^{-1} \in K[u]$ et par suite $u_{n+1} \in K[u]$. En vertu de

l'identité remarquable

$$a^m - b^m = (a-b)(a^{m-1} + a^{m-2}b + \cdots + ab^{m-2} + b^{m-1}),$$

valable dans tout anneau commutatif, on a alors

$$P'(u_{n+1}) - P'(u_n) = (u_{n+1} - u_n) \circ Q(u), \quad Q \in K[X].$$

Or, $u_{n+1} - u_n = P(u_n) \circ R(u)$, $R \in K[X]$ (on a en fait $R(u) = -P'_n(u_n)^{-1}$). Puisque $P(u_n) \in P(u)^{2^n} K[u]$ et que $P(u)$ est nilpotent, on en déduit que l'endomorphisme $u_{n+1} - u_n$ est nilpotent (cela fonctionne parce que $K[u]$ est commutatif!!), et donc $P'(u_{n+1}) - P'(u_n)$ est nilpotent. Comme $P'(u_n)$ est inversible dans $K[u]$ et que $P'(u_{n+1}) - P'(u_n)$ est un élément nilpotent de $K[u]$, $P'(u_{n+1})$ est inversible dans $K[u]$ par le lemme XV-4.16 (2).

Il reste à voir que $P(u_{n+1}) \in P(u)^{2^{n+1}} K[u]$. Remarquons maintenant que pour tous a, $b \in K[u]$, et tout $Q \in K[X]$, on a

$$Q(a+b) - Q(a) - Q'(a) \circ b \in b^2 K[u].$$

En effet, il suffit de le voir pour $Q = X^m$, $m \geqslant 0$. Mais alors, on a

$$(a+b)^m = a^m + ma^{m-1} \circ b + \frac{m(m-1)}{2} a^{m-2} \circ b^2 + \ldots,$$

d'où l'assertion, puisque $Q' = mX^{m-1}$ dans ce cas.

En prenant $a = u_n$, $b = -P(u_n) \circ P'(u_n)^{-1}$, $Q = P$, on obtient alors

$$P(u_{n+1}) = P\big(u_n - P(u_n) \circ P'(u_n)^{-1}\big) \in P(u_n)^2 K[u],$$

et comme $P(u_n) \in P(u)^{2^n} K[u]$, on a le résultat souhaité. Cela achève la récurrence, ainsi que la démonstration. □

3.16. Remarque. Ce qui précède donne une méthode de calcul de la décomposition de Dunford. Si de plus K est de caractéristique nulle, on n'est même pas obligé de connaître explicitement les valeurs propres de u. En effet, si λ est une racine de multiplicité m de χ_u, alors λ est une racine de multiplicité $m-1$ de χ'_u par le corollaire XV-6.28.

On a donc $\dfrac{\chi_u}{\operatorname{pgcd}(\chi_u, \chi'_u)} = P$. Or, ce pgcd peut se calculer grâce à l'algorithme d'Euclide, sans connaître la factorisation de χ_u. En particulier, la décomposition peut se trouver de manière totalement algorithmique, même le calcul de l'inverse de $P'(u_n)$.

Voici l'algorithme en détail lorsque K est de caractéristique nulle.

(1) On calcule $P = \dfrac{\chi_u}{\operatorname{pgcd}(\chi_u, \chi'_u)}$, et l'on pose $u_0 = u$.

(2) Si $P(u) = 0$, on s'arrête. On a $\delta = u$, $\nu = 0$.

Sinon, on détermine $U, V \in K[X]$ tels que $UP' + V\chi_u = 1$ avec l'algorithme d'Euclide. On a alors $P'(u_0)^{-1} = U(u)$.

On calcule ensuite $u_1 = u - P(u_0) \circ P'(u_0)^{-1}$.

(3) Supposons avoir calculé u_n et $P'(u_{n-1})^{-1}$ pour un $n \geqslant 1$. Si $P(u_n) = 0$, on s'arrête, et l'on pose $\delta = u_n$, $\nu = u - u_n$. Sinon, on calcule $P'(u_n)^{-1}$ comme suit :

(1) on calcule un entier $m \geqslant 1$ tel que $\big(P'(u_n) - P'(u_{n-1})\big)^m = 0$ (un tel entier existe, car $P'(u_n) - P'(u_{n-1})$ est nilpotent) ;

(2) on pose $a = P'(u_n) - P'(u_{n-1}), b = P'(u_{n-1})^{-1}$. On calcule $P'(u_n)$ par la formule

$$P'(u_n)^{-1} = \sum_{k=0}^{m-1} (-1)^k a^k \circ b^{k+1}.$$

Cela donne bien l'inverse de $P'(u_n)$. En effet, on a

$$P'(u_n) = a + b^{-1} = (\mathrm{Id}_E + a \circ b) \circ b^{-1}.$$

Comme $a^m = 0$ et que a et b commutent (puisque ce sont des polynômes en u), on a $(a \circ b)^m = a^m \circ b^m = 0$, et donc

$$(\mathrm{Id}_E + a \circ b)^{-1} = \sum_{k=0}^{m-1} (-1)^k a^k \circ b^k.$$

Ainsi,

$$P'(u_n)^{-1} = b \circ (\mathrm{Id}_E + a \circ b)^{-1} = \sum_{k=0}^{m-1} (-1)^k a^k \circ b^{k+1},$$

la dernière égalité découlant du fait que a et b commutent.

(4) On calcule ensuite u_{n+1} et l'on repart à l'étape précédente en remplaçant n par $n + 1$.

L'algorithme fonctionne bien entendu si l'on remplace χ_u par un polynôme annulateur P_0 divisant χ_u, et l'on aurait intérêt à prendre $P_0 = \mu_u$, mais le calcul de μ_u à partir de χ_u suppose que l'on sache trouver les diviseurs de χ_u, et donc le factoriser. En fait, il existe une formule générale pour calculer μ_u sans factoriser χ_u, mais elle n'est pas très pratique et très coûteuse en calculs (cf. corollaire XXXI-4.5).

La décomposition de Dunford est satisfaisante en ce qui concerne le calcul de puissances de matrices. En effet, comme déjà expliqué, elle le réduit au calcul des puissances quelconques d'un endomorphisme diagonalisable, ce qui devient simple après diagonalisation (sous réserve que l'on sache déterminer les valeurs propres), et au calcul d'un nombre fini de puissances d'une matrice nilpotente.

En revanche, la décomposition de Dunford ne répond pas complètement au problème de classification. Si $A \in \mathrm{M}_n(K)$, et si $A = D + N$ est la décomposition de Dunford de A, alors pour tout $P \in \mathrm{GL}_n(K)$, on a

$$PAP^{-1} = PDP^{-1} + PNP^{-1}.$$

On vérifie aisément que cette décomposition est la décomposition de Dunford de PAP^{-1}. En particulier, si $A = D + N$ et $A' = D' + N'$, alors A et A' sont conjuguées si, et seulement si, il existe une matrice $P \in \mathrm{GL}_n(K)$ telle que $D = PD'P^{-1}$ et $N = PN'P^{-1}$.

Il est facile de se convaincre que deux matrices diagonalisables sont conjuguées si, et seulement si, elles ont même polynôme caractéristique. En revanche, il n'est a priori pas évident de savoir lorsque deux matrices nilpotentes sont conjuguées. Remarquons que, même si l'on sait répondre à cette question, le problème de classification n'est pas pour autant résolu. En effet, il faut ensuite trouver une condition nécessaire et suffisante pour l'existence d'une matrice inversible qui conjugue **à la fois** D' en D et N' en N.

Pour pouvoir résoudre complètement le problème de classification, il faut raffiner la décomposition de Dunford. C'est le but de la section suivante.

4. Décomposition de Jordan

Commençons par définir les matrices qui interviendront de manière cruciale dans la suite.

4.1. Définition. Soit $r \geqslant 1$ un entier et soit $\lambda \in K$. La matrice

$$J_{r,\lambda} = \begin{pmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \in \mathrm{M}_r(K)$$

est appelée une *cellule de Jordan* de taille r . Un *bloc de Jordan* de taille r associé à λ est une matrice diagonale par blocs de la forme

$$\begin{pmatrix} J_{r_1,\lambda} & & \\ & \ddots & \\ & & J_{r_h,\lambda} \end{pmatrix} \in \mathrm{M}_r(K),$$

avec $1 \leqslant r_1 \leqslant \cdots \leqslant r_h \leqslant r$ et $r_1 + \cdots + r_h = r$. Remarquons que si $r = 1$, on a $J_{r,\lambda} = (\lambda)$.

Soit $u \in \mathscr{L}(E)$. Le but de cette section est de démontrer l'existence d'une base (sous certaines conditions), appelée *base de Jordan* , dans laquelle la matrice représentative est de la forme

$$\begin{pmatrix} J_{\lambda_1} & & \\ & \ddots & \\ & & J_{\lambda_r} \end{pmatrix} \in \mathrm{M}_n(K),$$

où $J_{\lambda_1}, \ldots, J_{\lambda_r}$ sont des blocs de Jordan associés à des éléments $\lambda_1, \ldots, \lambda_r$ de K deux à deux distincts. Une telle matrice sera appelé une *forme de Jordan* de u.

Remarquons que si une telle base existe, la matrice correspondante est triangulaire. En particulier, $\lambda_1, \ldots, \lambda_r \in K$ sont les valeurs propres de u et le polynôme χ_u est scindé.

Supposons maintenant que χ_u soit scindé, et soit $E = \bigoplus_{\lambda \in \mathrm{Sp}_K(u)} E_\lambda^{m_\lambda}$ la décomposition de E en somme directe de sous-espaces caractéristiques.

Admettons un instant que l'on ait montré l'existence d'une base de Jordan pour les endomorphismes nilpotents. D'après la proposition 3.7, le sous-espace $E_\lambda^{m_\lambda}$ est stable par u, et $u_\lambda = (u - \lambda \mathrm{Id}_E)_{E_\lambda^{m_\lambda}} \in \mathscr{L}(E_\lambda^{m_\lambda})$ est nilpotent. Soit $\mathbf{e}_\lambda$ est une base de Jordan de $E_\lambda^{m_\lambda}$ pour u_λ. Alors, la matrice représentative $\mathrm{Mat}(u_{E_\lambda^{m_\lambda}}; \mathbf{e}_\lambda)$ sera un bloc de Jordan associé à λ (car la matrice de l'endomorphisme identité dans n'importe quelle base est la matrice identité !). En recollant ces bases pour former une base de E, et puisque chaque sous-espace caractéristique est stable par u, la matrice représentative de u dans cette base sera diagonale par blocs, dont les blocs diagonaux seront les blocs de Jordan associés aux valeurs propres de u. En d'autres termes, on aura obtenu une base de Jordan associée à u.

Le cas crucial est donc le cas d'un endomorphisme nilpotent. On commence par un lemme.

4.2. Lemme. *Soit $u \in \mathscr{L}(E)$ un endomorphisme nilpotent, et soit $r \geqslant 1$ l'indice de nilpotence de u. Alors, on a*

$$\mathrm{Ker}(u^{i-1}) \subsetneq \mathrm{Ker}(u^i), \quad \textit{pour tout } i \in [\![1, r]\!]$$

et

$$\mathrm{Ker}(u^i) = \mathrm{Ker}(u^r) = E, \quad \textit{pour tout } i \geqslant r.$$

Démonstration. Soit $i \geqslant 1$ un entier. Si $i \geqslant r$, on a $u^i = 0$ et $\mathrm{Ker}(u^i) = E$. Supposons que $i \in [\![1, r]\!]$. Il est clair que $\mathrm{Ker}(u^{i-1}) \subset \mathrm{Ker}(u^i)$. Supposons que l'on ait égalité, et montrons que l'on a alors $\mathrm{Ker}(u^i) = \mathrm{Ker}(u^{i+1})$.

Comme précédemment, l'inclusion $\mathrm{Ker}(u^i) \subset \mathrm{Ker}(u^{i+1})$ est évidente. Soit maintenant $x \in E$ tel que $u^{i+1}(x) = 0$. Alors, $u^i\big(u(x)\big) = 0$, et par consé-

quent $u(x) \in \mathrm{Ker}(u^i) = \mathrm{Ker}(u^{i-1})$. On a ainsi

$$u^i(x) = u^{i-1}\big(u(x)\big) = u^{i-1}(0) = 0,$$

d'où $x \in \mathrm{Ker}(u^i)$. Par récurrence, on obtient

$$\mathrm{Ker}(u^{i-1}) = \mathrm{Ker}(u^i) = \mathrm{Ker}(u^{i+1}) = \cdots = \mathrm{Ker}(u^r) = E.$$

Mais alors $u^{i-1} = 0$, ce qui contredit le fait que r soit l'indice de nilpotence de u, puisque $i - 1 \leqslant r - 1 < r$. □

On peut maintenant démontrer la proposition suivante.

4.3. Proposition. *Tout endomorphisme nilpotent admet une base de Jordan.*

Démonstration. Soit $u \in \mathscr{L}(E)$, d'indice de nilpotence $r \geqslant 1$. Nous allons construire des sous-espaces $F_0, F_1, \ldots, F_r, F_{r+1}$ de E tels que

$$\mathrm{Ker}(u^i) = \mathrm{Ker}(u^{i-1}) \oplus F_i \quad \text{et} \quad u(F_i) \subset F_{i-1}, \qquad \text{pour tout } i \in [\![1, r+1]\!]$$

On procède par récurrence descendante. Posons $F_{r+1} = \{0\}$. Les propriétés requises sont clairement vérifiées pour $i = r+1$, même si F_r n'est pas encore construit, puisque $u(F_{r+1}) = \{0\}$.

Soit $i \in [\![1, r+1]\!]$. Supposons F_i construit, et construisons F_{i-1}. On a

$$\mathrm{Ker}(u^i) = \mathrm{Ker}(u^{i-1}) \oplus F_i.$$

Pour tout $x \in F_i$, on a

$$u^{i-1}\big(u(x)\big) = u^i(x) = 0,$$

car $F_i \subset \mathrm{Ker}(u^i)$. Ainsi, $u(F_i) \subset \mathrm{Ker}(u^{i-1})$.

Supposons tout d'abord que $i \geqslant 2$. Les sous-espaces $u(F_i)$ et $\mathrm{Ker}(u^{i-2})$ sont alors en somme directe. En effet, si $x = u(y) \in u(F_i)$ vérifie $u^{i-2}(x) = 0$, on a alors $y \in F_i \cap \mathrm{Ker}(u^{i-1}) = \{0\}$, et par conséquent $x = u(y) = 0$. On définit ensuite le sous-espace F_{i-1} comme étant un supplémentaire quelconque de $\mathrm{Ker}(u^{i-2})$ dans $\mathrm{Ker}(u^{i-1})$ contenant $u(F_i)$. Pour obtenir un tel supplémentaire, on colle une base de $\mathrm{Ker}(u^{i-2})$ et une base de $u(F_i)$. Cela fournit une famille libre, car ces deux sous-espaces sont en somme directe, et l'on complète en une base de $\mathrm{Ker}(u^{i-1})$. Les éléments de la base qui ne sont pas dans $\mathrm{Ker}(u^{i-2})$ engendrent un sous-espace qui est le supplémentaire recherché.

Si $i = 1$, puisque $\mathrm{Ker}(u^0) = \mathrm{Ker}(\mathrm{Id}_E) = \{0\}$, on a $F_1 = \mathrm{Ker}(u)$. On pose alors $F_0 = \{0\}$, et les propriétés requises sont aussi vérifiées.

Cela achève la récurrence descendante. Notons au passage que les sous-espaces $F_1, \ldots, F_r$ sont non nuls par le lemme 4.2. Remarquons maintenant

que la restriction de u à F_i est injective pour $i \geqslant 2$. En effet, si $x \in F_i$ vérifie $u(x) = 0$, alors

$$x \in F_i \cap \mathrm{Ker}(u) \subset F_i \cap \mathrm{Ker}(u^{i-1}) = \{0\}.$$

Remarquons aussi que l'on a

$$E = F_1 \oplus \cdots \oplus F_r.$$

On peut alors construire notre base de Jordan. On note

$$d_i = \dim_K \left(\mathrm{Ker}(u^i)\right) - \dim_K \left(\mathrm{Ker}(u^{i-1})\right) = \dim_K(F_i).$$

On va construire par récurrence descendante une base $(e_{i,1}, \ldots, e_{i,d_i})$ de F_i appropriée, pour tout $i \in [\![1, r]\!]$. On démarre, à cet effet, en choisissant une base $(e_{r,1}, \ldots, e_{r,d_r})$ de F_r. Soit $i \in [\![1, r]\!]$, et supposons avoir construit une base $(e_{i,1}, \ldots, e_{i,d_i})$ de F_i. Puisque la restriction de u à F_i est injective, la famille

$$e_{i-1,1} = u(e_{i,1}), \ \ldots, \ e_{i-1,d_i} = u(e_{i,d_i})$$

est une famille libre de F_{i-1}. On peut donc la compléter en une base

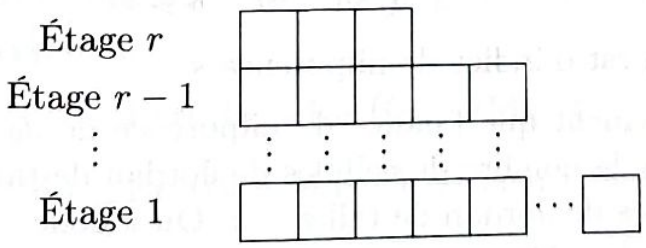
$$(e_{i-1,1}, \ \ldots, \ e_{i-1,d_{i-1}}) \quad \text{de } F_{i-1}.$$

Au bout du compte, on obtient une base $(e_{1,1}, \ \ldots, \ e_{1,d_1})$ de $F_1 = \mathrm{Ker}(u)$. La famille

$$(e_{1,1}, \ \ldots, \ e_{1,d_1}, \ \ldots, \ e_{r,1}, \ \ldots, \ e_{r,d_r})$$

est alors une base de E. Nous allons montrer qu'en ordonnant ces vecteurs de manière convenable, on obtient une base de Jordan. On les dispose en tableau de r étages de la façon suivante. Les étages sont numérotés de bas en haut, et sur l'étage i, on dispose la base de F_i.

On a alors un tableau de la forme



Par construction de la base précédente, les vecteurs sur l'étage 1 ont une image nulle par u (puisque $F_1 = \mathrm{Ker}(u)$), et pour tout $i \geqslant 2$, l'image d'un vecteur de l'étage i par u est le vecteur juste en dessous dans l'étage $i - 1$. Remarquons que les vecteurs d'une colonne fixée, ordonnés en partant du bas, engendrent un sous-espace stable par u, et E est la somme directe de ces sous-espaces. La matrice de la restriction de u à un de ces sous-espaces est une cellule de Jordan dont la taille est la hauteur de la colonne. En effet, soient $v_1, \ldots, v_s$ les vecteurs d'une colonne fixée. D'après ce qui précède,

on a

$$u(v_1) = 0 \quad \text{et} \quad u(v_i) = v_{i-1},\ i \in [\![2, s]\!].$$

Pour avoir une base de Jordan, on ordonne donc les vecteurs de la base comme suit : on part de la colonne la plus à droite, et l'on commence à numéroter les vecteurs de bas en haut. Une fois tout en haut, on redescend en base de la colonne voisine de gauche, et l'on recommence le procédé jusqu'à épuisement des colonnes. □

Étudions alors de plus près la structure d'une forme de Jordan de u.

Notation. Soit $u \in \mathscr{L}(E)$. Pour tout $\lambda \in \mathrm{Sp}_K(u)$, et tout $k \geqslant 1$, on pose

$$d_{k,\lambda} = \dim_K\left(\mathrm{Ker}\left((u - \lambda \mathrm{Id}_E)^k\right)\right) - \dim_K\left(\mathrm{Ker}\left((u - \lambda \mathrm{Id}_E)^{k-1}\right)\right).$$

4.4. Lemme. *Soit $u \in \mathscr{L}(E)$ un endomorphisme nilpotent, d'indice de nilpotence r, et soit*

$$J_0 = \begin{pmatrix} J_{r_1,0} & & \\ & \ddots & \\ & & J_{r_h,0} \end{pmatrix}$$

une forme de Jordan de u. Alors, $r_h = r$. En particulier, la taille d'une cellule de Jordan de J_0 est inférieure ou égale à r, et pour tout $k \in [\![1, r]\!]$, le nombre de cellules de Jordan de taille $\geqslant k$ de J_0 est égal à $d_{k,0}$.

En particulier, le nombre de cellules de Jordan de taille k de J_0 est égal à

$$d_{k,0} - d_{k+1,0}.$$

Démonstration. Soit $J_{s,0}$ une cellule de Jordan de taille s. Un simple calcul matriciel montre que l'on a

$$\dim_K\left(\mathrm{Ker}(J_{s,0}^k)\right) = \begin{cases} k, & \text{si} \quad k \in [\![1, s-1]\!] \\ s, & \text{si} \quad k \geqslant s. \end{cases}$$

En particulier, $J_{s,0}$ est d'indice de nilpotence s.

On en déduit facilement que l'indice de nilpotence de J_0 est r_h, et donc que $r = r_h$. Soit n_k le nombre de cellules de Jordan de taille k et soit m_k le nombre de cellules de Jordan de taille $\geqslant k$. On a donc

$$m_k = n_k + \cdots + n_r.$$

Si $J_{s,0}$ est une cellule de Jordan de taille $s \in [\![1, k-1]\!]$, on a $J_{s,0}^k = 0$, et donc $\mathrm{Ker}(J_{s,0}^k)$ est de dimension s. Si $s \in [\![k, r]\!]$, $\mathrm{Ker}(J_{s,0}^k)$ est de dimension k. Alors, $\dim_K\left(\mathrm{Ker}(J_0^k)\right)$ est égale à

$$n_1 + 2n_2 + \cdots + (k-1)n_{k-1} + k(n_k + \cdots + n_r),$$

c'est-à-dire à

$$\dim_K\left(\mathrm{Ker}(J_0^k)\right) = n_1 + 2n_2 + \cdots + (k-1)n_{k-1} + km_k.$$

On a donc
$$\begin{aligned} d_{k,0} &= (k-1)n_{k-1} + km_k - (k-1)m_{k-1} \\ &= m_k + (k-1)(n_{k-1} + m_k - m_{k-1}) \\ &= m_k, \end{aligned}$$
car $n_{k-1} + m_k = n_{k-1} + n_k + \cdots + n_r = m_{k-1}$. Autrement dit, $m_k = d_{k,0}$. On a donc $n_k = m_k - m_{k+1} = d_{k,0} - d_{k+1,0}$. On en déduit alors le lemme. □

Nous pouvons maintenant énoncer le théorème principal de cette section.

4.5. Théorème (Réduction de Jordan). *Soit $u \in \mathscr{L}(E)$ tel que χ_u soit scindé. Alors, μ_u est aussi scindé, et u admet une base de Jordan. Écrivons*
$$\mu_u = \prod_{\lambda \in \mathrm{Sp}_K(u)} (X-\lambda)^{r_\lambda}.$$
Alors, la taille maximale d'une cellule de Jordan du bloc J_λ est r_λ. De plus, pour tout $k \in [\![1, r_\lambda]\!]$, le nombre de cellules de Jordan de type $J_{k,\lambda}$ dans le bloc J_λ est
$$d_{k,\lambda} - d_{k+1,\lambda}.$$
Enfin, u admet une forme de Jordan, unique une fois choisie une numérotation des valeurs propres, et deux endomorphismes $u, u' \in \mathscr{L}(E)$ dont les polynômes caractéristiques sont scindés sont semblables si, et seulement si, ils ont même forme de Jordan.

Démonstration. Puisque $\mu_u \mid \chi_u$, μ_u est aussi scindé. De plus, χ_u et μ_u ont les mêmes racines par la proposition 1.19. Les racines de μ_u sont donc les valeurs propres de u. On a vu plus haut que pour obtenir une base de Jordan pour u, il suffisait d'obtenir une base de Jordan de la restriction de $u - \lambda \mathrm{Id}_E$ à $E_\lambda^{m_\lambda}$, qui est nilpotente, et cela pour tout $\lambda \in \mathrm{Sp}_K(u)$, ce qui est possible d'après la proposition 4.3. On a donc l'existence d'une base de Jordan pour u.

Soit $\lambda \in \mathrm{Sp}_K(u)$ et soit $u_\lambda = (u - \lambda \mathrm{Id}_E)_{E_\lambda^{m_\lambda}}$. Remarquons que pour tout entier $i \in [\![0, m_\lambda]\!]$, on a
$$\mathrm{Ker}(u_\lambda^i) = \mathrm{Ker}\big((u - \lambda \mathrm{Id}_E)^i\big).$$
L'inclusion $\mathrm{Ker}(u_\lambda^i) \subset \mathrm{Ker}\big((u - \lambda \mathrm{Id}_E)^i\big)$ est claire, puisque
$$u_\lambda^i = \big((u - \lambda \mathrm{Id}_E)_{E_\lambda^{m_\lambda}}\big)^i = \big((u - \lambda \mathrm{Id}_E)^i\big)_{E_\lambda^{m_\lambda}}.$$
Supposons maintenant que $x \in \mathrm{Ker}\big((u - \lambda \mathrm{Id}_E)^i\big)$. Alors, on a
$$\mathrm{Ker}\big((u - \lambda \mathrm{Id}_E)^i\big) \subset E_\lambda^{m_\lambda},$$
car $i \leqslant m_\lambda$. Mais alors on a
$$0 = (u - \lambda \mathrm{Id}_E)^i(x) = (u - \lambda \mathrm{Id}_E)^i_{E_\lambda^{m_\lambda}}(x) = u_\lambda^i(x),$$
d'où l'inclusion manquante.

Par la proposition 3.7, u_λ est nilpotent, d'indice de nilpotence r_λ. La deuxième partie provient alors du lemme 4.4 et de la remarque ci-dessus. Cela montre également que la structure de J_λ est entièrement déterminée par u, et donc u admet une unique forme de Jordan, une fois une numérotation des valeurs propres choisies. Si u et u' ont même forme de Jordan, alors ils sont semblables puisqu'ils ont même matrice représentative dans des bases de E bien choisies. Enfin, si u et u' sont semblables, ils ont même polynôme minimal et de plus les sous-espaces $\mathrm{Ker}\big((u-\lambda\mathrm{Id}_E)^i\big)$ et $\mathrm{Ker}\big((u'-\lambda\mathrm{Id}_E)^i\big)$ sont isomorphes pour tout $i \geqslant 0$, donc de même dimension.

En effet, s'il existe $v \in \mathrm{GL}(E)$ tel que $u = v \circ u' \circ v^{-1}$, alors

$$(u-\lambda\mathrm{Id}_E)^i = \big(v\circ(u'-\lambda\mathrm{Id}_E)\circ v^{-1}\big)^i = v\circ(u'-\lambda\mathrm{Id}_E)^i\circ v^{-1},$$

et cela pour tout $i \geqslant 0$.

Il en résulte que pour tout $x \in E$, on a $x \in \mathrm{Ker}\big((u-\lambda\mathrm{Id}_E)^i\big)$ si, et seulement si, $v(x) \in \mathrm{Ker}\big((u'-\lambda\mathrm{Id}_E)^i\big)$. Ainsi, v induit un isomorphisme entre les sous-espaces $\mathrm{Ker}\big((u-\lambda\mathrm{Id}_E)^i\big)$ et $\mathrm{Ker}\big((u'-\lambda\mathrm{Id}_E)^i\big)$.

Il en résulte que u et u' ont même forme de Jordan. □

Résumons la méthode pratique pour obtenir la forme de Jordan d'un endomorphisme $u \in \mathscr{L}(E)$, ainsi qu'une base de Jordan, que l'on peut extraire des démonstrations précédentes. On commence par calculer μ_u, et l'on écrit

$$\mu_u = \prod_{\lambda\in\mathrm{Sp}_K(u)} (X-\lambda)^{r_\lambda}.$$

Pour ce faire, on peut procéder ainsi. On calcule χ_u, et l'on obtient ainsi la multiplicité m_λ de $\lambda \in \mathrm{Sp}_K(u)$. Alors, r_λ est le plus petit entier $r \geqslant 1$ tel que $\dim_K\big(\mathrm{Ker}\big((u-\lambda\mathrm{Id}_E)^r\big)\big) = m_\lambda$.

(1) Si l'on veut seulement la forme de Jordan de u : pour tout $\lambda \in \mathrm{Sp}_K(u)$ et tout $k \in [\![1, r_\lambda]\!]$, on calcule

$$d_{k,\lambda} = \dim_K\big(\mathrm{Ker}\big((u-\lambda\mathrm{Id}_E)^k\big)\big) - \dim_K\big(\mathrm{Ker}\big((u-\lambda\mathrm{Id}_E)^{k-1}\big)\big).$$

Le nombre de cellules de Jordan de taille k dans le bloc de Jordan J_λ est alors égal à $d_{k,\lambda} - d_{k+1,\lambda}$.

Pour lire ces nombres facilement, on peut construire un tableau à r_λ étages tel que pour tout entier $k \in [\![1, r_\lambda]\!]$, l'étage k ait $d_{k,\lambda}$ cases. Alors, $d_{k,\lambda} - d_{k+1,\lambda}$ est le nombre de colonnes de hauteur k (remarquer que $d_{r_\lambda+1} = 0$). On en déduit alors la structure de J_λ et l'on colle tous les blocs dans une matrice diagonale par blocs pour obtenir la forme de Jordan de u.

(2) Si l'on veut calculer une base de Jordan : pour tout $\lambda \in \mathrm{Sp}_K(u)$, on construit un tableau à r_λ étages formé de vecteurs de E comme suit. On choisit tout d'abord des vecteurs formant une base d'un supplémentaire de $\mathrm{Ker}\big((u-\lambda\mathrm{Id}_E)^{r_\lambda-1}\big)$ dans $\mathrm{Ker}\big((u-\lambda\mathrm{Id}_E)^{r_\lambda}\big)$, et on les dispose en ligne au dernier étage. On applique ensuite $u-\lambda\mathrm{Id}_E$ à ces vecteurs, et l'on complète cette nouvelle famille en une base d'un supplémentaire de $\mathrm{Ker}\big((u-\lambda\mathrm{Id}_E)^{r_\lambda-2}\big)$ dans $\mathrm{Ker}\big((u-\lambda\mathrm{Id}_E)^{r_\lambda-1}\big)$, que l'on dispose à l'étage du dessous. On répète le procédé jusqu'à obtenir une base de $\mathrm{Ker}(u-\lambda\mathrm{Id}_E)$ que l'on dispose à l'étage 1. On renumérote ensuite les vecteurs en commençant par parcourir la colonne la plus à droite de bas en haut, puis en recommençant le procédé avec la colonne voisine de gauche, jusqu'à épuisement des colonnes. On recolle alors toutes les familles de vecteurs obtenues ainsi pour chaque valeur propre, et l'on obtient une base de Jordan de u.

4.6. Remarque. Tout ce qui précède a bien sûr un équivalent matriciel.

4.7. Exemple. Soit $M \in \mathrm{M}_6(\mathbb{C})$ la matrice

$$M = \begin{pmatrix} 3 & -1 & -3 & 0 & -1 & 3 \\ 0 & 2 & -3 & 0 & 0 & 3 \\ 0 & 0 & -1 & 0 & 0 & 3 \\ -1 & 1 & -3 & 2 & 1 & 4 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 2 \end{pmatrix}.$$

On vérifie que $\chi_M = (X+1)(X-2)^5$. Alors, $\mathrm{Ker}(M+\mathrm{I}_6)$ est de dimension 1 (car -1 est une valeur propre simple), engendré par

$$e_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

On a donc $r_{-1} = 1$ et le bloc de Jordan correspondant est (-1). On vérifie que l'on a

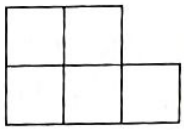
$$\mathrm{Ker}(M-2\mathrm{I}_6) = \{\mathbf{x} \in \mathbb{C}^6 \mid x_3 = x_6 = 0,\ x_5 = x_1 - x_2\}$$

et

$$\mathrm{Ker}\big((M-2\mathrm{I}_6)^2\big) = \{\mathbf{x} \in \mathbb{C}^6 \mid x_3 = x_6\}.$$

On a donc $\dim_K\big(\mathrm{Ker}(M-2\mathrm{I}_6)\big) = 3$ et $\dim_K\big(\mathrm{Ker}(M-2\mathrm{I}_6)^2\big) = 5$. Puisque la multiplicité de 2 est égale à 5, il est inutile d'aller plus loin. Par consé-

quent, $r_2 = 2$, $d_{1,-2} = 3$, $d_{2,-2} = 5 - 3 = 2$, et l'on a le tableau suivant pour la valeur propre 2 :



La numérotation d'une base de Jordan pour le bloc J_2 sera donc

e_6	e_4	
e_5	e_3	e_2

On a donc une cellule de taille 1, et deux cellules de taille 2, et le bloc correspondant est donc

$$\begin{pmatrix} 2 & & & & & \\ & 2 & 1 & & & \\ & 0 & 2 & & & \\ & & & 2 & 1 & \\ & & & 0 & 2 & \end{pmatrix}.$$

La forme de Jordan de M est ainsi

$$\begin{pmatrix} -1 & & & & & \\ & 2 & & & & \\ & & 2 & 1 & & \\ & & 0 & 2 & & \\ & & & & 2 & 1 \\ & & & & 0 & 2 \end{pmatrix}.$$

Calculons une base de Jordan. Les vecteurs e_6 et e_4 doivent être une base d'un supplémentaire de $\mathrm{Ker}\big((M - 2\mathrm{I}_6)\big)$ dans $\mathrm{Ker}\big((M - 2\mathrm{I}_6)^2\big)$. On vérifie que les vecteurs

$$e_6 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \quad \text{et} \quad e_4 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \\ -1 \\ 1 \end{pmatrix}$$

conviennent[4]. On calcule ensuite

$$e_5 = (M - 2\mathrm{I}_6)e_6 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{et} \quad e_3 = (M - 2\mathrm{I}_6)e_4 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix},$$

4. Attention ! Il ne suffit pas de prendre deux vecteurs libres dans le complémentaire de $\mathrm{Ker}\big((M - 2\mathrm{I}_6)\big)$ dans $\mathrm{Ker}\big((M - 2\mathrm{I}_6)^2\big)$!

que l'on complète en une base de $\mathrm{Ker}(M - 2\mathrm{I}_6)$ en prenant

$$e_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}.$$

Une fois que l'on a trouvé une forme de Jordan d'une matrice A, il est facile de calculer ses puissances successives. En effet, tout se réduit à savoir calculer les puissances successives d'une cellule de Jordan $J_{r,\lambda}$. Cela est particulièrement aisé.

4.8. Lemme. *Pour tout $r \geqslant 1$, tout $m \geqslant 0$, et tout $\lambda \in K$, on a*

$$J_{r,\lambda}^m = \begin{pmatrix} \lambda^m & m\lambda^{m-1} & \binom{m}{2}\lambda^{m-2} & \cdots & \binom{m}{r-1}\lambda^{m-(r-1)} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \binom{m}{2}\lambda^{m-2} \\ & & & \ddots & m\lambda^{m-1} \\ & & & & \lambda^m \end{pmatrix},$$

où l'on convient que $\binom{m}{k}\lambda^{m-k} = 0$ si $k > m$.

Démonstration. On remarque que $J_{r,\lambda} = \lambda \mathrm{I}_n + J_{r,0}$. On a alors

$$J_{r,\lambda}^m = \sum_{k=0}^{m} \binom{m}{k} \lambda^{m-k} J_{r,0}^k.$$

Or, les matrices $J_{r,0}^k$ sont faciles à calculer : on a $J_{r,0}^0 = \mathrm{I}_r$, et si $k \geqslant 1$, on décale la « diagonale » de 1 d'un cran vers la droite, jusqu'à obtenir la matrice nulle pour $k \geqslant r$. Le résultat s'ensuit. □

Nous achevons cette section par quelques remarques générales.

4.9. Remarques

(1) Tous ces résultats supposent que le polynôme caractéristique de la matrice A considérée est scindé. Ce n'est pas une restriction très lourde. En effet, si K est un corps quelconque, on a les résultats suivants :
(i) il existe une extension L/K dans laquelle χ_A est scindé (cf. proposition XXVI-1.6) ;

(ii) si deux matrices de $\mathrm{M}_n(K)$ sont conjuguées par une matrice inversible $Q \in \mathrm{GL}_n(L)$, alors elles sont en fait déjà conjuguées par une matrice inversible $P \in \mathrm{GL}_n(K)$.
Le lecteur sera reportera au lemme 5.10 si $K = \mathbb{R}$ et $L = \mathbb{C}$, et au corollaire XXXI-4.9 pour le cas général.

Fort de ces deux résultats, on voit que la réduction de Jordan permet malgré tout de déterminer si deux matrices de $\mathrm{M}_n(K)$ sont conjuguées. En ce qui concerne le calcul des puissances de A, il suffit d'utiliser la réduction de Jordan ou la décomposition de Dunford dans L. Le résultat final sera quand même une matrice de $\mathrm{M}_n(K)$, puisque A est à coefficients dans K.

(2) Il existe une méthode plus simple pour calculer les puissances successives d'une matrice $A \in \mathrm{M}_n(K)$ possédant un polynôme caractéristique scindé, lorsque K est de caractéristique nulle. Soit $P \in K[X]$ un polynôme annulateur scindé (par exemple, χ_A ou μ_A). Si $d = \deg(P) \geqslant 1$, alors, par division euclidienne, on a

$$X^m = PQ_m + a_m^{(d-1)} X^{d-1} + \cdots + a_m^{(1)} X + a_m^{(0)} \qquad (*)$$

pour tout $m \geqslant 0$.
En évaluant en A, on obtient

$$A^m = a_m^{(d-1)} A^{d-1} + \cdots + a_m^{(1)} A + a_m^{(0)} \mathrm{I}_n,$$

car $P(A) = 0$. L'idée est donc de calculer les coefficients du reste de la division euclidienne de X^m par P sans en avoir à calculer le quotient. Pour cela, on procède de la manière suivante.

Pour chaque racine $\alpha \in K$ de P, de multiplicité s_α, on remarque que α est une racine de la dérivée k-ième de PQ_n, pour $k \in [\![0, s_\alpha - 1]\!]$. On dérive alors $s_\alpha - 1$ fois l'équation (*), pour obtenir s_α équations dans lesquelles on évalue en A. En faisant cela pour chaque racine, on obtient un système de d équations linéaires à d inconnues, que l'on résout pour trouver les coefficients voulus.

Si P n'est pas scindé sur K, on fait les calculs dans une extension L/K dans laquelle P se scinde.

La justification de cette méthode est basée sur le résultat suivant, dont la démonstration est laissée au lecteur : soit K un corps de caractéristique nulle. Soient r, $d \geqslant 1$, soient $\alpha_1, \ldots, \alpha_r \in K$ deux à deux distincts, et soient $s_1, \ldots, s_r \geqslant 1$ tels que $s_1 + \cdots + s_r = d$. Alors, pour toute famille $(\beta_{ij})_{i \in [\![1,r]\!],\, j \in [\![0, s_i - 1]\!]}$ d'éléments de K, il existe un unique polynôme $R \in K[X]$ de degré $\leqslant d - 1$ tel que

$$R^{(j)}(\alpha_i) = \beta_{ij}, \quad \text{pour tous } i \in [\![1, r]\!],\ j \in [\![0, s_i - 1]\!].$$

Remarquons que ce résultat n'est plus vrai si K est de caractéristique non nulle.

En effet, si $K = \mathbb{F}_2$, alors $P_1 = X^2 + 1$ et $P_2 = 1$ sont deux polynômes de degré $\leqslant 2$ tels que $P_i(0) = 1$, $P_i'(0) = P_i''(0) = 0$.

Nous finissons ce long chapitre en étudiant l'exponentielle de matrices.

5. Exponentielle de matrices

Dans tout ce qui suit, le corps K désignera $\mathbb{R}$ ou $\mathbb{C}$, cela afin d'assurer la convergence absolue de la série exponentielle introduite ci-après.

On définit d'abord une norme sur $\mathrm{M}_n(K)$ de la manière suivante.

Si $A = (a_{ij}) \in \mathrm{M}_n(K)$, on pose

$$|A| = \max_{1 \leqslant i \leqslant n} \sum_{j=1}^{n} |a_{ij}|.$$

Elle vérifie de plus

$$|AB| \leqslant |A||B|,$$

pour A, $B \in \mathrm{M}_n(K)$.

5.1. Lemme. *Soit $A \in \mathrm{M}_n(K)$. La série $\displaystyle\sum_{m \geqslant 0} \frac{A^m}{m!}$ est convergente.*

Démonstration. Puisque $(\mathrm{M}_n(K), |\cdot|)$ est un K-espace vectoriel normé de dimension finie, il est complet. Il suffit donc de montrer que la série est absolument convergente. Mais, on a

$$\left|\frac{A^m}{m!}\right| \leqslant \frac{|A|^m}{m!}.$$

La série de terme général $\dfrac{|A|^m}{m!}$ étant convergente (elle converge vers $e^{|A|}$), les théorèmes de comparaison sur les séries à termes positifs permettent de conclure. □

5.2. Définition. Pour $A \in \mathrm{M}_n(K)$, on définit l'exponentielle de A par

$$e^A = \sum_{m \geqslant 0} \frac{A^m}{m!}.$$

Voici maintenant quelques propriétés élémentaires de l'exponentielle.

5.3. Proposition. *On a les propriétés suivantes :*

(1) *pour tout $A \in E$, A et e^A commutent ;*

(2) *si A et B commutent, on a $e^{A+B} = e^A e^B$;*

(3) *pour tout $A \in \mathrm{M}_n(K)$, e^A est inversible et $(e^A)^{-1} = e^{-A}$;*

(4) *pour tout $P \in \mathrm{GL}_n(K)$ et pour tout $A \in \mathrm{M}_n(K)$, on a*

$$e^{PAP^{-1}} = Pe^A P^{-1}.$$

Démonstration. Remarquons que les applications

$$\begin{array}{ccc} \mathrm{M}_n(K) & \longrightarrow & \mathrm{M}_n(K) \\ M & \longmapsto & AM \end{array} \qquad \text{et} \qquad \begin{array}{ccc} \mathrm{M}_n(K) & \longrightarrow & \mathrm{M}_n(K) \\ M & \longmapsto & MA \end{array}$$

sont linéaires, donc continues, puisque nous sommes en dimension finie.

On a donc $Ae^A = A\Big(\sum\limits_{m\geqslant 0} \dfrac{A^m}{m!}\Big) = \sum\limits_{m\geqslant 0} \Big(A\dfrac{A^m}{m!}\Big)$, et par conséquent

$$Ae^A = \sum_{m\geqslant 0} \Big(\frac{A^m}{m!}A\Big) = \Big(\sum_{m\geqslant 0} \frac{A^m}{m!}\Big)A = e^A A,$$

d'où (1). Si maintenant A et B commutent, on peut appliquer le théorème sur le produit de séries absolument convergentes et utiliser la formule du binôme de Newton pour conclure. Les détails sont laissés au lecteur. On obtient ainsi (2). Le point (3) est alors immédiat par le point précédent, en remarquant que $e^0 = \mathrm{I}_n$.

Il reste à démontrer (4). L'application

$$\begin{array}{ccc} \mathrm{M}_n(K) & \longrightarrow & \mathrm{M}_n(K) \\ M & \longmapsto & PMP^{-1} \end{array}$$

est linéaire continue, comme composée de deux applications linéaires continues (à savoir la multiplication à gauche par P, et la multiplication à droite par P^{-1}). On raisonne alors comme précédemment, en remarquant que

$$(PAP^{-1})^n = PA^nP^{-1}, \quad \text{pour tout } n \geqslant 0.$$

Cela achève la démonstration. □

Une application majeure de l'exponentielle de matrices est la résolution de systèmes d'équations linéaires à coefficients constants. Nous n'en parlerons pas ici, et renvoyons le lecteur à son cours d'analyse.

La difficulté première concernant l'exponentielle de matrices est son calcul pratique, sa définition même rendant le dit calcul a priori inaccessible, sauf dans quelques cas particuliers.

Par exemple si A est diagonalisable, alors le point (4) de la proposition

précédente permet de calculer e^A sans effort. En effet,

$$\text{si } A = P\begin{pmatrix}\lambda_1 & & \\ & \ddots & \\ & & \lambda_n\end{pmatrix}P^{-1}, \text{ alors } e^A = P\begin{pmatrix}e^{\lambda_1} & & \\ & \ddots & \\ & & e^{\lambda_n}\end{pmatrix}P^{-1}.$$

Par ailleurs, à l'autre extrême, lorsque A est nilpotente, le calcul de e^A se ramène au calcul d'une somme finie.

La décomposition de Dunford permet de réduire le calcul à ces deux cas faciles. En effet, par la proposition 5.3 (4), si $A = D+N$ est la décomposition de Dunford de A, comme D et N commutent, on a

$$e^A = e^D e^N = e^D\Big(\mathrm{I}_n + N + \frac{N^2}{2!} + \cdots + \frac{N^{r-1}}{r!}\Big),$$

où r est l'indice de nilpotence de N.

Pour achever le calcul, il suffit alors de diagonaliser la matrice D afin de calculer e^D, ce qui présuppose le calcul des valeurs propres de A (qui sont aussi celles de D) et de leurs multiplicités. Cela dit, puisque l'on doit a priori calculer les valeurs propres, il vaut mieux utiliser le résultat suivant, qui réduit le calcul de l'exponentielle aux calcul des projecteurs spectraux, ce que l'on sait faire de manière algorithmique (cf. remarque 3.14).

5.4. Proposition. *Soit $A \in \mathrm{M}_n(K)$. Pour tout $\lambda \in \mathrm{Sp}_K(A)$, on note Π_λ la matrice du projecteur spectral associé à λ dans la base canonique.*

Alors, on a

$$e^A = \sum_\lambda e^\lambda \sum_{k=0}^{n-1} \frac{(A-\lambda \mathrm{I}_n)^k}{k!}\Pi_\lambda.$$

En particulier, e^A est un polynôme en A.

Démonstration. Soit u l'endomorphisme de K^n dont la matrice représentative dans la base canonique est A. Soit $p_\lambda \in \mathscr{L}(K^n)$ le projecteur spectral associé à λ.

En utilisant les définitions, on vérifie facilement que $p_\lambda^2 = p_\lambda$, et $p_\lambda p_\mu = 0$ dès que $\lambda \neq \mu$. On a donc $\Pi_\lambda^2 = \Pi_\lambda$, et $\Pi_\lambda \Pi_\mu = 0$ dès que $\lambda \neq \mu$.

De même, $\sum_{\lambda \in \mathrm{Sp}_K(u)} p_\lambda = \mathrm{Id}_{K^n}$, par définition, et donc

$$\sum_{\lambda \in \mathrm{Sp}_K(A)} \Pi_\lambda = \mathrm{I}_n.$$

Rappelons également que la décomposition de Dunford de A est donnée par $A = D + N$, avec $D = \sum_{\lambda \in \mathrm{Sp}_K(A)} \lambda \Pi_\lambda$ et $N = A - D$.

On commence par calculer $e^{\lambda\Pi_\lambda}$. Pour tout $m \geqslant 1$, on a $(\lambda\Pi_\lambda)^m = \lambda^m\Pi_\lambda$, car $\Pi_\lambda^2 = \Pi_\lambda$. Alors, on a

$$e^{\lambda\Pi_\lambda} = \mathrm{I}_n + \Big(\sum_{m\geqslant 1} \frac{\lambda^m}{m!}\Big)\Pi_\lambda = \mathrm{I}_n + (e^\lambda - 1)\Pi_\lambda.$$

Puisque l'on a $\Pi_\lambda\Pi_\mu = 0$ si $\lambda \neq \mu$, les Π_λ commutent, donc

$$e^D = \prod_{\lambda\in \mathrm{Sp}_K(A)} \big(\mathrm{I}_n + (e^\lambda - 1)\Pi_\lambda\big),$$

et puisque $\Pi_\lambda\Pi_\mu = 0$ pour $\lambda \neq \mu$, on obtient alors

$$e^D = \mathrm{I}_n + \sum_{\lambda\in \mathrm{Sp}_K(A)} (e^\lambda - 1)\Pi_\lambda = \sum_{\lambda\in \mathrm{Sp}_K(A)} e^\lambda\Pi_\lambda,$$

car $\sum\limits_{\lambda\in \mathrm{Sp}_K(A)} \Pi_\lambda = 1$.

On aurait pu le voir en faisant un changement de base qui diagonalise la matrice $\sum\limits_{\lambda\in \mathrm{Sp}_K(A)} \lambda\Pi_\lambda$ et utiliser le point (4) de la proposition 5.3.

Puisque $N \in \mathrm{M}_n(K)$ est nilpotente, on a $N^n = 0$. Comme Π_λ et N sont des polynômes en A par le lemme 3.9 et le théorème 3.10, ils commutent. On a donc

$$e^A = e^D e^N = \Big(\sum_{\lambda\in \mathrm{Sp}_K(A)} e^\lambda\Pi_\lambda\Big)\Big(\sum_{k=0}^{n-1} \frac{N^k}{k!}\Big) = \sum_{\lambda\in \mathrm{Sp}_K(A)} e^\lambda \sum_{k=0}^{n-1} \frac{N^k}{k!}\Pi_\lambda.$$

Il reste donc à montrer que pour tout $k \geqslant 0$, on a $N^k\Pi_\lambda = (A - \lambda\mathrm{I}_n)^k\Pi_\lambda$.

Le cas $k = 0$ étant trivial, on suppose $k \geqslant 1$. Mais alors $\Pi_\lambda = \Pi_\lambda^k$, et puisque N et Π_λ commutent, on a alors $N^k\Pi_\lambda = N^k\Pi_\lambda^k = (N\Pi_\lambda)^k$.

Mais, $N = A - \sum\limits_{\lambda\in \mathrm{Sp}_K(A)} \lambda\Pi_\lambda$, et on obtient alors $N\Pi_\lambda = A\Pi_\lambda - \lambda\Pi_\lambda$, en utilisant les relations entre les matrices Π_λ. Alors,

$$(N\Pi_\lambda)^k = \big((A - \lambda\mathrm{I}_n)\Pi_\lambda\big)^k = (A - \lambda\mathrm{I}_n)^k\Pi_\lambda^k,$$

car A et Π_λ commutent, et l'on conclut en utilisant l'égalité $\Pi_\lambda^k = \Pi_\lambda$.

Le dernier point s'obtient en utilisant le fait que les matrices Π_λ sont des polynômes en A. □

5.5. Remarque. On peut remplacer n dans la somme précédente par le maximum des indices de nilpotence des matrices $(A - \lambda\mathrm{I}_n)\Pi_\lambda$.

Notons que la réduction de Jordan permet de réduire le calcul de l'exponentielle au calcul de l'exponentielle d'une cellule de Jordan. Celui-ci est très

simple. L'exponentielle de matrices étant en particulier utile pour résoudre des systèmes d'équations différentielles linéaires, on va plutôt donner la formule pour $e^{tJ_{\lambda,r}}$, et l'on laissera le lecteur spécialiser au cas $t = 1$.

5.6. Lemme. *Pour tous t, $\lambda \in K$, et tout $r \geqslant 1$, on a*

$$e^{tJ_{r,\lambda}} = e^{\lambda t} = \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^m}{m!} \\ & \ddots & \ddots & \ddots & \vdots \\ & & \ddots & \ddots & \frac{t^2}{2!} \\ & & & \ddots & t \\ & & & & 1 \end{pmatrix},$$

Démonstration. On a $J_{r,\lambda} = \lambda \mathrm{I}_r + J_{r,0}$, d'où

$$e^{tJ_{r,\lambda}} = e^{\lambda t \mathrm{I}_r} e^{tJ_{r,0}} = e^{\lambda t} e^{tJ_{r,0}}.$$

On peut donc supposer sans perte de généralité que $\lambda = 0$. Il reste donc à calculer $\dfrac{t^m J_{r,0}^m}{m!}$, ce qui est immédiat, et donne le résultat attendu. □

On va maintenant s'intéresser à la question suivante : quelle est l'image de l'application exponentielle ? Pour cela, on considère le cas réel et le cas complexe séparément. Le cas complexe est relativement aisé à traiter à l'aide de la théorie des séries formelles.

5.7. Théorème. *L'application*

$$\begin{aligned} M_n(\mathbb{C}) &\longrightarrow \mathrm{GL}_n(\mathbb{C}) \\ A &\longmapsto e^A \end{aligned}$$

est surjective. Plus précisément, pour tout $A \in \mathrm{GL}_n(\mathbb{C})$, il existe $P \in \mathbb{C}[X]$ tel que $A = e^{P(A)}$.

Démonstration. Soit N une matrice nilpotente. On va montrer que $\mathrm{I}_n + N$ est l'exponentielle d'un polynôme en N.

Posons $E = e^X \in \mathbb{C}[[X]]$ et $L = \ln(1+X) \in \mathbb{C}[[X]]$. Remarquons que pour toute matrice $M \in \mathrm{M}_n(\mathbb{C})$ nilpotente, on a $E(M) = e^M$.

D'après la proposition XXIII-3.2, on a

$$E \circ L = 1 + X \in \mathbb{C}[X].$$

En appliquant la proposition XXIII-2.10 à $A = \mathbb{C}$, $B = \mathbb{C}[N]$ et $b = N$, on

obtient alors $E\big(L(N)\big) = \mathrm{I}_n + N$, soit

$$e^{L(N)} = \mathrm{I}_n + N.$$

Comme $L(N)$ est un polynôme en N d'après cette même proposition, on a le résultat voulu.

Soit maintenant $A \in \mathrm{GL}_n(\mathbb{C})$ une matrice inversible. On écrit grâce à la décomposition spectrale $A = D + N$, avec $D = \sum_{\lambda \in \mathrm{Sp}_{\mathbb{C}}(A)} \lambda \Pi_\lambda$.

Comme A est inversible, ses valeurs propres sont non nulles. Puisque l'image de l'exponentielle complexe est $\mathbb{C}^\times$, on peut écrire $\lambda = e^{\alpha_\lambda}$, $\alpha_\lambda \in \mathbb{C}$.

Posons

$$D' = \sum_{\lambda \in \mathrm{Sp}_{\mathbb{C}}(A)} \alpha_\lambda \Pi_\lambda.$$

En raisonnant comme dans la proposition 5.4, on voit que

$$e^{D'} = \sum_{\lambda \in \mathrm{Sp}_{\mathbb{C}}(A)} e^{\alpha_\lambda} \Pi_\lambda = D.$$

De plus, D' est un polynôme en A, car chaque Π_λ est un polynôme en A.

On a d'une part $A = D(\mathrm{I}_n + D^{-1}N)$. Or, D et N commutent (ce sont des polynômes en A), donc D^{-1} et N aussi, et $D^{-1}N$ est nilpotente, puisque alors $(D^{-1}N)^s = (D^{-1})^s N^s = 0$, où s est l'indice de nilpotence de N. Par le premier point, $\mathrm{I}_n + D^{-1}N = e^L$, où L est un polynôme en $D^{-1}N$.

D'autre part, si $\chi_D(X) = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0$, on a $a_0 \neq 0$, puisque D est inversible et que a_0 est le déterminant de D au signe près. L'égalité $\chi_D(D) = 0$ entraîne alors que $D^{-1} = -a_0^{-1}(D^{n-1} + \cdots + a_1 \mathrm{I}_n)$, et donc D^{-1} est un polynôme en D, donc en A, et il en est alors de même de $D^{-1}N$, et donc de L.

Mais alors, L et D' étant des polynômes en A, ils commutent et donc

$$A = D + N = D(\mathrm{I}_n + D^{-1}N) = e^{D'}e^L = e^{D'+L},$$

ce qui achève la démonstration. □

5.8. Corollaire. *Pour toute matrice A inversible, et tout $n \geqslant 1$, il existe un polynôme $Q \in \mathbb{C}[X]$ tel que $A = Q(A)^n$.*

Démonstration. Par le théorème précédent, il existe $P \in \mathbb{C}[X]$ tel que

$$A = e^{P(A)} = e^{\frac{P(A)}{n} + \cdots + \frac{P(A)}{n}} = \Big(e^{\frac{P(A)}{n}}\Big)^n.$$

Or, $e^{\frac{P(A)}{n}}$ est un polynôme en $P(A)$, donc en A, d'où le résultat. □

L'exponentielle n'est bien sûr pas injective, puisque les matrices 0 et $2i\pi \mathrm{I}_n$ sont toutes deux envoyées sur I_n.

On passe au cas où $K = \mathbb{R}$.

Remarquons tout de suite que l'exponentielle n'est plus surjective. En effet, si $A \in \mathrm{M}_n(\mathbb{R})$, en trigonalisant A dans $\mathbb{C}$, on voit que si $\lambda_1, \ldots, \lambda_n$ sont les valeurs propres **complexes** de A (non nécessairement distinctes), alors les valeurs propres complexes de e^A sont $e^{\lambda_1}, \ldots, e^{\lambda_n}$. En particulier,

$$\det e^A = e^{\mathrm{tr}A} > 0,$$

et l'image de l'exponentielle est donc contenue dans le sous-groupe $\mathrm{GL}_n^+(\mathbb{R})$ des matrices réelles inversibles de déterminant strictement positif. On montrera plus tard que l'image de l'exponentielle n'est pas $\mathrm{GL}_n^+(\mathbb{R})$ tout entier.

Mais pour l'instant, donnons quelques résultats partiels qui nous seront utiles par la suite.

5.9. Proposition. *Si* $\mathrm{Sp}_{\mathbb{C}}(B) \subset \mathbb{R}^{+\times}$, *alors* $B \in e^{\mathrm{M}_n(\mathbb{R})}$.

Démonstration. En effet, l'exponentielle réelle réalise une surjection sur de $\mathbb{R}$ sur $\mathbb{R}^{+\times}$, et la démonstration du théorème 5.7 reste alors valable. □

Continuons par un petit lemme très classique.

5.10. Lemme. *Soient deux matrice réelles* U *et* V *semblables sur* $\mathbb{C}$. *Alors, les matrices* U *et* V *sont semblables sur* $\mathbb{R}$.

Démonstration. Soit $M = X + iY \in \mathrm{GL}_n(\mathbb{C})$ telle que $M^{-1}UM = V$. Ainsi, on a $UM = MV$, puis $UX = XV$ et $UY = YV$, en identifiant partie réelle et partie imaginaire. On en tire facilement que

$$U(X + \lambda Y) = (X + \lambda Y)V,$$

pour tout $\lambda \in \mathbb{C}$. Or, $\det(X + tY)$ est un polynôme non nul de $\mathbb{R}[t]$. En effet, s'il était nul, il serait aussi nul comme polynôme de $\mathbb{C}[t]$. En particulier, le déterminant $\det(X + iY)$ serait nul, ce qui n'est pas le cas, puisque M est inversible. On peut donc choisir $\lambda \in \mathbb{R}$ tel que $\det(X + \lambda Y)$ soit non nul. Pour un tel réel λ, $X + \lambda Y$ est une matrice réelle inversible, et l'on a alors

$$(X + \lambda Y)^{-1}U(X + \lambda Y) = V.$$

Cela achève la démonstration. □

Introduisons quelques notations.

Si $A = (a_{ij}) \in \mathrm{M}_n(\mathbb{C})$ est une matrice complexe, on pose $\overline{A} = (\overline{a}_{ij})$, ainsi que

$$\mathrm{Re}\,A = \frac{A + \overline{A}}{2}, \quad \mathrm{Im}\,A = \frac{A - \overline{A}}{2i} \quad \text{et} \quad A^* = \begin{pmatrix} \mathrm{Re}\,A & -\mathrm{Im}\,A \\ \mathrm{Im}\,A & \mathrm{Re}\,A \end{pmatrix}.$$

On remarque que ces trois dernières matrices sont réelles.

On pose $M = \begin{pmatrix} i\mathrm{I}_n & \mathrm{I}_n \\ \mathrm{I}_n & i\mathrm{I}_n \end{pmatrix}$.

On vérifie sans peine que M est inversible, d'inverse

$$M^{-1} = \frac{1}{2}\begin{pmatrix} -i\mathrm{I}_n & \mathrm{I}_n \\ \mathrm{I}_n & -i\mathrm{I}_n \end{pmatrix}.$$

On a alors aisément $\begin{pmatrix} A & 0 \\ 0 & \overline{A} \end{pmatrix} = M^{-1}A^*M$.

Ces considérations vont nous permettre de montrer la proposition suivante.

5.11. Proposition. *Soit B une matrice réelle inversible. On suppose que B est semblable sur $\mathbb{C}$ à une matrice de la forme $\begin{pmatrix} A & 0 \\ 0 & \overline{A} \end{pmatrix}$. Alors, $B \in e^{\mathrm{M}_n(\mathbb{R})}$.*

Démonstration. Par hypothèse, il existe $P \in \mathrm{GL}_n(\mathbb{C})$ et $A \in \mathrm{M}_n(\mathbb{C})$ telles que

$$PBP^{-1} = \begin{pmatrix} A & 0 \\ 0 & \overline{A} \end{pmatrix}.$$

On a $\det B = \det(A)\det(\overline{A})$, et comme $\det B \neq 0$ par hypothèse, on a également $\det A \neq 0$, donc A est inversible. Par le théorème 5.7, on a

$$A = e^U, \quad U \in M_n(\mathbb{C}).$$

L'application

$$\begin{array}{ccc} \mathrm{M}_n(\mathbb{C}) & \longrightarrow & \mathrm{M}_n(\mathbb{C}) \\ U & \longmapsto & \overline{U} \end{array}$$

étant clairement continue, on a alors $\overline{e^U} = e^{\overline{U}}$, pour toute matrice complexe U.

On a donc $\overline{A} = e^{\overline{U}}$, et donc

$$\begin{pmatrix} A & 0 \\ 0 & \overline{A} \end{pmatrix} = \begin{pmatrix} e^U & 0 \\ 0 & e^{\overline{U}} \end{pmatrix} = e^{\begin{pmatrix} U & 0 \\ 0 & \overline{U} \end{pmatrix}},$$

d'où

$$\begin{aligned} (MP)B(MP)^{-1} &= M\begin{pmatrix} A & 0 \\ 0 & \overline{A} \end{pmatrix}M^{-1} = Me^{\begin{pmatrix} U & 0 \\ 0 & \overline{U} \end{pmatrix}}M^{-1} \\ &= e^{M\begin{pmatrix} U & 0 \\ 0 & \overline{U} \end{pmatrix}M^{-1}} = e^{U^*}. \end{aligned}$$

Les matrices réelles B et e^{U^*} sont ainsi semblables sur $\mathbb{C}$, donc sur $\mathbb{R}$ par le lemme précédent. La conclusion vient alors facilement. □

5.12. Remarque. Cela entraîne en particulier que toute matrice de la forme
$$\begin{pmatrix} A & 0 \\ 0 & A \end{pmatrix}, \quad A \in \mathrm{GL}_n(\mathbb{R})$$
est l'exponentielle d'une matrice réelle.

5.13. Corollaire. *Soit $B \in \mathrm{GL}_n(\mathbb{R})$ une matrice sans valeurs propres réelles. Alors, $B \in e^{\mathrm{M}_n(\mathbb{R})}$.*

Démonstration. Soit Λ l'ensemble des valeurs propres de B de partie imaginaire > 0. Le lemme de décomposition des noyaux nous donne alors :
$$\mathbb{C}^n = \bigoplus_{\lambda\in\Lambda} \mathrm{Ker}\big((B-\lambda \mathrm{I}_n)^{m_\lambda}\big) \oplus \bigoplus_{\lambda\in\Lambda} \mathrm{Ker}\big((B-\overline{\lambda}\mathrm{I}_n)^{m_\lambda}\big).$$
On pose $F = \bigoplus_{\lambda\in\Lambda} \mathrm{Ker}\big((B-\lambda \mathrm{I}_n)^{m_\lambda}\big)$ et $H = \bigoplus_{\lambda\in\Lambda} \mathrm{Ker}\big((B-\overline{\lambda}\mathrm{I}_n)^{m_\lambda}\big)$. Ce sont des sous-espaces stables par B.

Montrons que $H = \{\overline{X} \mid X \in F\}$.

On a
$$\begin{aligned} (B-\lambda \mathrm{I}_n)^{m_\lambda} X = 0 &\iff \overline{(B-\lambda \mathrm{I}_n)^{m_\lambda} X} = 0 \\ &\iff (\overline{B}-\overline{\lambda}\mathrm{I}_n)^{m_\lambda}\overline{X} = 0 \\ &\iff (B-\overline{\lambda}\mathrm{I}_n)^{m_\lambda}\overline{X} = 0, \end{aligned}$$
car B est coefficients réels. Cela permet de conclure facilement.

En particulier si $(e_1,\dots,e_r)$ est une base de F, alors $(\overline{e_1},\dots,\overline{e_r})$ est une base de H. Remarquons enfin que $B\overline{e_i} = \overline{Be_i}$, car à nouveau B est à coefficients réels.

Par conséquent, la matrice de l'endomorphisme de $\mathbb{C}^n$ représenté par B dans la base $(e_1,\dots,e_r,\overline{e_1},\dots,\overline{e_r})$ est de la forme
$$\begin{pmatrix} A & 0 \\ 0 & \overline{A} \end{pmatrix}.$$
Autrement dit, B est semblable sur $\mathbb{C}$ à une matrice de cette forme. On utilise alors la proposition précédente. □

Avant de continuer, remarquons que l'exponentielle ne se comporte en général pas très bien par restriction à un sous-espace stable. Plus précisément, soit $A \in \mathrm{M}_n(\mathbb{R})$, et soit F un sous-espace de $\mathbb{R}^n$ stable par A. L'endomorphisme de matrice A se restreint alors en un endomorphisme de F. Soit alors A_F la matrice de cette restriction dans une base de F. Alors, si A est l'exponentielle d'une matrice réelle, il n'en est pas forcément de même de la matrice A_F.

En effet, la matrice $A = -\mathrm{I}_2$ est l'exponentielle d'une matrice réelle par la remarque précédente. Sa restriction à $\mathbb{R}\varepsilon_1$ (où ε_1 est le premier vecteur de la base canonique de $\mathbb{R}^2$) est la matrice (-1), qui n'est pas dans l'image de l'exponentielle réelle.

Le lemme suivant nous donne une condition suffisante pour que cela devienne vrai.

5.14. Lemme. *Soit $A \in \mathrm{GL}_n(\mathbb{R})$, et soit $F = \mathrm{Ker}\big(P(A)\big)$, avec $P \in \mathbb{R}[X]$. On suppose que $A \in e^{M_n(\mathbb{R})}$.*

Si u est l'endomorphisme de $\mathbb{R}^n$ représenté par A, alors n'importe quelle matrice représentative de u_F est l'exponentielle d'une matrice réelle.

Démonstration. Écrivons $A = e^M$, $M \in \mathrm{M}_n(\mathbb{R})$. Alors, M et $e^M = A$ commutent, donc M et $P(A)$ commutent. On en déduit facilement que le sous-espace $F = \mathrm{Ker}\big(P(A)\big)$ est stable par M.

Soit G un supplémentaire de F. On construit une nouvelle base de $\mathbb{R}^n$ en accolant une base de F et une base de G. Soit Q la matrice de passage correspondante. On a alors

$$Q^{-1}AQ = \begin{pmatrix} A_F & * \\ 0 & * \end{pmatrix} \quad \text{et} \quad Q^{-1}MQ = \begin{pmatrix} M_F & * \\ 0 & * \end{pmatrix},$$

où A_F et M_F désignent respectivement les matrices des restrictions de l'endomorphisme correspondant à A et M dans la base de F choisie. Mais alors, on a

$$Q^{-1}e^MQ = e^{Q^{-1}MQ} = \begin{pmatrix} e^{M_F} & * \\ 0 & * \end{pmatrix}.$$

Comme $A = e^M$, on en déduit que $A_F = e^{M_F}$, d'où le résultat. □

Nous sommes presque au bout de nos peines.

Soit $A \in \mathrm{GL}_n(\mathbb{R})$. On décompose χ_A sous la forme

$$\chi_A = PP_-P_+,$$

où P est le produit des facteurs irréductibles de degré 2, P_+ (resp. P_-) est le produit des facteurs $(X - \lambda)$, $\lambda > 0$ (resp. $\lambda < 0$), et on pose

$$F = \mathrm{Ker}\big(P(A)\big) \quad \text{et} \quad F_\pm = \mathrm{Ker}\big(P_\pm(A)\big).$$

Le lemme de décomposition des noyaux nous assure alors que

$$\mathbb{R}^n = F \oplus F^+ \oplus F^-.$$

Soit $u \in \mathscr{L}(\mathbb{R}^n)$ l'endomorphisme représenté par A.

On a alors le résultat suivant.

5.15. Lemme. *Avec les notations précédentes, A est l'exponentielle d'une matrice réelle si, et seulement si, la matrice représentative de la restriction de u à F_- dans une base arbitraire l'est.*

Démonstration. Le sens direct est une application du lemme 5.14. Montrons la réciproque.

Choisissons des bases de F, F_+ et F_-, et construisons une base de $\mathbb{R}^n$ par recollement des bases. Si Q est la matrice de changement de base, on a

$$Q^{-1}AQ = \begin{pmatrix} A_F & & \\ & A_+ & \\ & & A_- \end{pmatrix},$$

où A_F, A_+ et A_- désignent respectivement les matrices des endomorphismes de F, F_+ et F_- induits par u dans les bases choisies, et cela par stabilité par A de ces trois sous-espaces. Par le lemme 1.23, les polynômes caractéristiques de A_F, A_+ et A_- sont respectivement P, P_+ et P_-.

Remarquons que A_F n'a pas de valeurs propres réelles et que A_+ n'a que des valeurs propres réelles > 0. D'après la proposition 5.9, il existe une matrice M_+ telle que $A_+ = e^{M_+}$. Par le corollaire 5.13, il existe une matrice M telle que $A_F = e^M$. Par hypothèse, il existe également une matrice M_- telle que $A_- = e^{M_-}$.

On a alors
$$Q^{-1}AQ = \begin{pmatrix} e^M & & \\ & e^{M_+} & \\ & & e^{M_-} \end{pmatrix} = e^B,$$
où $B = \begin{pmatrix} M & & \\ & M_+ & \\ & & M_- \end{pmatrix}$. On a ainsi $A = e^{QMQ^{-1}}$, ce qui achève la démonstration. □

5.16. Corollaire. *Soit $B \in \mathrm{GL}_n(\mathbb{R})$ une matrice sans valeurs propres réelles strictement négatives. Alors, $B \in e^{\mathrm{M}_n(\mathbb{R})}$.*

5.17. Corollaire. *Si $n \geqslant 2$, l'image de $\mathrm{M}_n(\mathbb{R})$ par l'exponentielle de matrices n'est pas $\mathrm{GL}_n^+(\mathbb{R})$ tout entier.*

Démonstration. On considère la matrice diagonale

$$A = \begin{pmatrix} -1 & & \\ & -2 & \\ & & \mathrm{I}_{n-2} \end{pmatrix}.$$

Clairement, $\det(A) > 0$. Soit ε_1 le premier vecteur de la base canonique. Si A était une exponentielle, alors sa restriction à $\mathrm{Ker}(A + \mathrm{I}_n) = \mathbb{R}^n\varepsilon_1$, soit (-1), le serait aussi par le lemme 5.14, ce qui est absurde car $-1 < 0$. □

Mais quelle est donc cette fichue image, alors ? Eh bien, chers lecteurs (enfin, ceux qui restent...), piaffant d'impatience, voici enfin la réponse.

5.18. Théorème. *On a* $e^{\mathrm{M}_n(\mathbb{R})} = \{M^2 \mid M \in \mathrm{GL}_n(\mathbb{R})\}$.

Démonstration. Si $A = e^M$, alors $A = \left(e^{\frac{M}{2}}\right)^2$. Réciproquement, on suppose que $A = B^2$, $B \in \mathrm{GL}_n(\mathbb{R})$.

On va proposer deux démonstrations du fait que $A \in e^{\mathrm{M}_n(\mathbb{R})}$.

Première démonstration. Comme B et A commutent, le sous-espace F_- associé à A est stable par B. En raisonnant exactement comme dans le lemme 5.15, on obtient $A_- = B_-^2$, avec des notations évidentes.

Les valeurs propres de A_- sont donc exactement les carrés de valeurs propres de B_- (trigonaliser B_- pour le voir). Puisque A_- n'a que des valeurs propres réelles strictement négatives, A_- est inversible, donc B_- aussi, et les valeurs propres de B_- sont toutes imaginaires pures non nulles. Par le corollaire 5.13, $B_- = e^C, C \in \mathrm{M}_n(\mathbb{R})$, et alors $A_- = (e^C)^2 = e^{2C}$. Par le lemme 5.15, A est l'exponentielle d'une matrice réelle.

Seconde démonstration. D'après le théorème 5.7, il existe $P \in \mathbb{C}[X]$ tel que $B = e^{P(B)}$. Comme B est à coefficients réels, on a

$$B = \overline{B} = \overline{e^{P(B)}} = e^{\overline{P(B)}} = e^{\overline{P}(B)}.$$

Puisque $P(B)$ et $\overline{P}(B)$ commutent, on a alors

$$A = B^2 = e^{P(B)}e^{\overline{P}(B)} = e^{P(B)+\overline{P}(B)} = e^{Q(B)},$$

avec $Q = P + \overline{P} \in \mathbb{R}[X]$. Comme $Q(B) \in \mathrm{M}_n(\mathbb{R})$, on obtient le résultat voulu. □

Terminons ce chapitre par un corollaire immédiat du théorème précédent.

5.19. Corollaire. *Soit* $A \in \mathrm{GL}_n(\mathbb{R})$. *Alors, A est un carré de* $\mathrm{GL}_n(\mathbb{R})$ *si, et seulement si, c'est une puissance n-ième pour tout* $n \geqslant 2$.

6. Exercices

Dans ce qui suit, $K_n[X]$ désigne l'espace vectoriel des polynômes de $K[X]$ de degré $\leqslant n$.

E1. Soit E l'espace vectoriel des suites $(u_n)_{n\geqslant 0}$ à valeurs dans K, et soit

$$\begin{array}{rccc} s : & E & \longrightarrow & E \\ & (u_n)_{n\geqslant 0} & \longmapsto & s\big((u_n)_{n\geqslant 0}\big) = (0, u_0, u_1, \ldots). \end{array}$$

Montrer que $\mathrm{Ann}(s) = (0)$.

E2. Commentez cette démonstration du théorème de Cayley-Hamilton : soit $M \in \mathrm{M}_n(K)$. Alors,

$$\chi_M(M) = \det(M\mathrm{I}_n - M) = \det(0) = 0.$$

E3. Soient $u_1, \ldots, u_n \in \mathscr{L}(E)$ des endomorphismes commutant deux à deux.

a. On suppose que $u_1, \ldots, u_n$ sont diagonalisables/trigonalisables. Montrer qu'ils sont diagonalisables/trigonalisables dans une même base.

b. En déduire que si un endomorphisme $u \in \mathscr{L}(E)$ possède une décomposition de Dunford, alors χ_u est scindé sur K.

E4. Soit $E = \mathbb{R}_3[X]$. On considère l'application u de E dans E ainsi définie : à tout polynôme P de E on associe le reste de la division de XP par $X^4 - 1$.

a. Montrer que u est un endomorphisme.

b. Quelle est la matrice de u dans la base canonique de E ?

c. Déterminer le polynôme caractéristique de u.

d. Déterminer le polynôme minimal de u.

e. L'endomorphisme u est-il diagonalisable ?

f. Montrer que $X^4 - 1$ se factorise dans $\mathbb{R}[X]$ en un produit de trois polynômes premiers entre eux deux à deux.

g. En déduire une décomposition de E en somme directe de trois sous-espaces stables par u, de dimensions respectives 1, 1 et 2. Décrire ces sous-espaces et la restriction de u à chacun d'eux en donnant par exemple sa matrice dans une base convenable. Donner une réduction diagonale par blocs associée et la matrice de passage correspondante.

E5. Soit $B \in \mathbb{R}[X]$. On considère l'application $f : \mathbb{R}[X] \longrightarrow \mathbb{R}[X]$ qui à P associe le reste de la division euclidienne de P par B.

a. Vérifier que f est un endomorphisme de $\mathbb{R}[X]$ et que $\mathbb{R}_n[X]$ est stable par f pour tout $n \geqslant 0$.

On note f_n l'endomorphisme de $\mathbb{R}_n[X]$ induit par f.

b. Pour $B = X^2 + X + 1$, écrire la matrice de f_3 dans la base canonique.

On revient au cas général.

c. Quels sont les valeurs propres et vecteurs propres de f_n (distinguer les cas $\deg(B) \geqslant n$ et $\deg(B) > n$) ?

d. L'endomorphisme f_n est-il diagonalisable ? Si oui, donner une base de vecteurs propres. Si non, f est-il trigonalisable ?

e. Mêmes questions avec f définie par $f(P) = XP'$.

E6. Sans utiliser la forme de Jordan, dire parmi les matrices suivantes, lesquelles sont semblables.

$$M_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad M_2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad M_3 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$M_4 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad M_5 = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad M_6 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix},$$

$$M_7 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad M_8 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad M_9 = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix},$$

$$M_{10} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

E7. Soit E un K-espace vectoriel de dimension finie $n \geqslant 2$. Montrer qu'un endomorphisme de rang 1 est diagonalisable si, et seulement si, sa trace est non nulle.

Indication. Montrer que 0 est valeur propre, et calculer sa multiplicité.

E8. Soit E un K-espace vectoriel de dimension finie, et soit $p \in \mathscr{L}(E)$ vérifiant $p^2 = p$. Comparer $\mathrm{rg}(p)$ et $\mathrm{tr}(p)$ (attention à la caractéristique !).

E9. Soit $M \in \mathrm{M}_n(\mathbb{C})$.

a. On suppose que M est annulée par $X(X-1)(X-2)$. Calculer M^m en fonction de I_n, M et M^2 pour tout $n \geqslant 3$.

b. Même question si M est annulée par $(X-1)^2(X-2)$.

E10. Soit E un $\mathbb{R}$-espace vectoriel de dimension $n \geqslant 1$, et soit u un endomorphisme de E vérifiant $u^2 + \mathrm{Id}_E = 0$. On se propose de montrer de

plusieurs manières que n est pair, et de déterminer tous les endomorphismes u vérifiant cette relation.

a. Utiliser le déterminant pour montrer que n est pair.

b. Soit M une matrice représentative de u dans une base de E fixée.

(i) Montrer que les valeurs propres complexes éventuelles de M sont i et $-i$.

(ii) Montrer de deux manières différentes le résultat suivant :
si $P \in \mathbb{R}[X]$, et si $z \in \mathbb{C}$ est une racine de P de multiplicité m, alors $\overline{z}$ est une racine de P de multiplicité m,
▷ en utilisant la caractérisation de la multiplicité en termes de dérivées successives ;
▷ en montrant que l'application

$$\begin{array}{ccc} \mathbb{C}[X] & \longrightarrow & \mathbb{C}[X] \\ P = \sum_{i=0}^{d} a_i X^i & \longmapsto & \overline{P} = \sum_{i=0}^{d} \overline{a}_i X^i \end{array}$$

est un morphisme d'anneaux, et en remarquant que $P \in \mathbb{R}[X]$ si, et seulement si, $P = \overline{P}$.

c. En déduire que $\chi_u = (X^2+1)^m$, $m \geqslant 1$. Retrouver le fait que n est pair.

d. (i) Montrer que la loi externe

$$\begin{array}{ccc} \mathbb{C} \times E & \longrightarrow & E \\ (a+bi, v) & \longmapsto & (a+bi)\cdot v = av + bu(v) \end{array}$$

définit une structure de $\mathbb{C}$-espace vectoriel sur E qui prolonge sa structure de $\mathbb{R}$-espace vectoriel.

(ii) Montrer que si $\mathbf{e} = (e_1, \ldots, e_m)$ est une $\mathbb{C}$-base de E, alors

$$\mathbf{e}' = (e_1, i\cdot e_1, \ldots, e_m, i\cdot e_m)$$

est une $\mathbb{R}$-base de E. Retrouver encore une fois que n est pair.

(iii) Calculer la matrice représentative de u dans la $\mathbb{R}$-base $\mathbf{e}'$ (remarquer que $i\cdot e_j = u(e_j)$ par définition).

(iv) En déduire que $u^2 + \mathrm{Id}_E = 0$ si, et seulement si, il existe une $\mathbb{R}$-base de E dans laquelle la matrice représentative de u est de la forme

$$\begin{pmatrix} 0 & -1 & & & \\ 1 & 0 & & & \\ & & \ddots & & \\ & & & 0 & -1 \\ & & & 1 & 0 \end{pmatrix}.$$

E11. Soit $M \in \mathrm{M}_2(\mathbb{Z})$. On suppose qu'il existe $n \geqslant 1$ tel que $M^n = \mathrm{I}_2$. Montrer que $M^{12} = \mathrm{I}_2$.

Indication. Quelles sont les valeurs propres dans $\mathbb{C}$? Calculer la trace.

E12. Soit u un endomorphisme d'un K-espace vectoriel E. On suppose que u est annulé par deux polynômes de $K[X]$ premiers entre eux. Que dire de E et u ?

E13. Que peut-on dire d'un endomorphisme u d'un $\mathbb{R}$-espace vectoriel de dimension finie annulé par $1 - X^3$ et $X^2 - 2X + 1$?

E14. Soit u un endomorphisme d'un K-espace vectoriel E de dimension finie.

a. On suppose que u est inversible. Montrer qu'il existe $Q \in K[X]$ tel que $u^{-1} = Q(u)$ (penser à Cayley-Hamilton).

b. Soit $P \in K[X]$. Montrer que $P(u)$ est inversible si, et seulement si, P et μ_u sont premiers entre eux.

E15. Vrai ou faux ?

Soit E un K-espace vectoriel de dimension finie. Soient u, $u' \in \mathscr{L}(E)$, et soient λ, $\lambda' \in K$.

a. Si $P \in K[X]$ et si λ est valeur propre de u, alors $P(\lambda)$ est valeur propre de $P(u)$.

b. Si λ est valeur propre de u et si λ' est valeur propre de u', alors $\lambda + \lambda'$ est valeur propre de $u + u'$.

c. Si λ est valeur propre de u et si λ' est valeur propre de u', alors $\lambda\lambda'$ est valeur propre de $u \circ u'$.

d. Si u est diagonalisable (resp. trigonalisable), alors u^2 est diagonalisable (resp. trigonalisable).

e. Si u^2 est diagonalisable (resp. trigonalisable), alors u est diagonalisable (resp. trigonalisable).

E16. Soit E un espace euclidien. Déterminer toutes les isométries de E qui soient diagonalisables.

E17. Soit p un nombre premier, soit E un $\mathbb{F}_p$-espace vectoriel de dimension finie et soit $u \in \mathscr{L}(E)$. Montrer que u est diagonalisable si, et seulement si, $u^p = u$.

Pour le lecteur connaissant la théorie des corps finis, généraliser le résultat dans le cas où $\mathbb{F}_p$ est remplacé par un corps fini.

E18. Soit E un K-espace vectoriel, et soit $u \in \mathscr{L}(E)$.

Soient $P_1, \dots, P_r \in K[X]$ des polynômes premiers entre eux deux à deux. On suppose que $P = P_1 \cdots P_r$ annule u, si bien que

$$E = \operatorname{Ker}\big(P_1(u)\big) \oplus \cdots \oplus \operatorname{Ker}\big(P_r(u)\big).$$

Pour $i \in [\![1, r]\!]$, on note $\pi_i \in \mathscr{L}(E)$ le projecteur sur $\operatorname{Ker}\big(P_i(u)\big)$ parallèlement à $\bigoplus_{j \neq i} \operatorname{Ker}\big(P_j(u)\big)$.

En s'inspirant de la démonstration du lemme 3.9, montrer que les projecteurs $\pi_1, \dots, \pi_r$ sont des polynômes en u.

E19. Soit E un K-espace vectoriel de dimension $n \geqslant 1$, et soit $u \in \mathscr{L}(E)$. On dit que u est *simple* si les seuls sous-espaces de E stables par u sont $\{0\}$ et E.

On veut montrer que u est simple si, et seulement si, χ_u est irréductible.

a. On suppose que χ_u est irréductible. Montrer que u est simple.
Indication. Si F est un sous-espace stable par u non nul, que vaut le polynôme χ_{u_F} ?

On suppose dorénavant que u est simple.

b. Soit $v \in E \setminus \{0\}$. Montrer que la famille $\big(v, u(v), \dots, u^{n-1}(v)\big)$ est une base de E. En déduire que $\mu_u = \chi_u$.

c. On suppose que $\chi_u = PQ$, avec $P, Q \in K[X]$. Montrer que l'on a

$$\operatorname{Ker}\big(P(u)\big) = E \quad \text{ou} \quad \operatorname{Ker}(Q(u)) = E.$$

d. En utilisant les questions précédentes, montrer que $\chi_u \mid P$ ou $\chi_u \mid Q$, et conclure.

E20. Soit E un K-espace vectoriel de dimension finie, et soit $u \in \mathscr{L}(E)$. On s'intéresse ici au problème suivant : tout sous-espace de E stable par u admet-il un supplémentaire stable par u ?

a. Soit $E = K^2$, et soit $u \in \mathscr{L}(E)$ l'endomorphisme dont la matrice représentative dans la base canonique est

$$M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Soit ε_1 le premier vecteur de la base canonique de K^2. Montrer que le sous-espace $F = K\varepsilon_1$ est l'unique droite de E est stable par u, et en déduire que F n'admet pas de supplémentaire stable par u.

On dit que $u \in \mathscr{L}(E)$ est *semi-simple* si tout sous-espace stable par u admet un supplémentaire stable par u.

b. On suppose que $\mu_u = \pi$, où $\pi \in K[X]$ un polynôme irréductible unitaire. On veut montrer que u est semi-simple. On pose $L = K[X]/(\pi)$.

(i) Montrer que l'application

$$\begin{aligned} L \times E &\longrightarrow E \\ (\overline{P}, x) &\longmapsto \overline{P} \bullet x = P(u)(x) \end{aligned}$$

est bien définie, et munit E d'une structure de L-espace vectoriel qui prolonge sa structure de K-espace vectoriel.

(ii) Montrer qu'un sous-espace F de E est stable par u si, et seulement si, F est un sous-espace du L-espace vectoriel E.

(iii) Soit F un sous-espace stable par u, et soit G un supplémentaire de F dans le L-espace vectoriel E. Montrer que G est un supplémentaire de F stable par u, et conclure.

On veut démontrer que u est semi-simple si, et seulement si, μ_u est un produit de polynômes irréductibles unitaires deux à deux distincts.

c. On suppose que $\mu_u = \pi_1 \cdots \pi_r$, où $\pi_1, \ldots, \pi_r \in K[X]$ sont irréductibles unitaires deux à deux distincts.

Pour $i \in [\![1, r]\!]$, on pose $F_i = \mathrm{Ker}\big(\pi_i(u)\big)$.

(i) Justifier que $\mu_{u_{F_i}} = \pi_i$.

(ii) Soit F un sous-espace de E stable par u. Montrer que l'on a

$$F = (F \cap F_1) \oplus \cdots \oplus (F \cap F_r).$$

(iii) Utiliser les questions précédentes pour montrer que u est semi-simple.

d. Soit $u \in \mathscr{L}(E)$ un endomorphisme semi-simple.

On suppose que $\mu_u = P^2 Q$, où $P, Q \in K[X]$ et P est non constant. Soit $F = \mathrm{Ker}\big(P(u)\big)$.

(i) Justifier que F admet un supplémentaire G stable par u.

(ii) Montrer que pour tout $x \in G$, on a $(PQ)(u)(x) \in F$. En déduire que $(PQ)(u)(x) = 0$ pour tout $x \in G$.

(iii) Montrer que PQ annule u et en déduire une contradiction.

e. Conclure.

f. On suppose que χ_u est scindé. Montrer u est semi-simple si, et seulement si, il est diagonalisable.

E21. Soit E un K-espace vectoriel de dimension $n \geqslant 1$, et soit $u \in \mathscr{L}(E)$ un endomorphisme de E. Soit $\pi \in K[X]$ un polynôme irréductible unitaire.

Si $\pi = X^d + a_{d-1}X^{d-1} + \cdots + a_0$, on pose

$$C_\pi = \begin{pmatrix} 0 & 0 & & 0 & -a_0 \\ 1 & 0 & & & -a_1 \\ 0 & 1 & & & -a_2 \\ & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & -a_{d-1} \end{pmatrix} \in \mathrm{M}_d(K).$$

Soit $u \in \mathscr{L}(E)$ tel que $\mu_u = \pi$. On pose $L = K[X]/(\pi)$.

a. Soit $(f_1, \ldots, f_r)$ une L-base de E pour la structure de L-vectoriel définie dans la question **b** (i) de l'exercice précédent. Montrer que la famille

$$\big(u^i(f_j)\big)_{i \in [\![0,d-1]\!],\, j \in [\![1,r]\!]}$$

est une K-base de E, où $d = \deg(\pi)$.

b. En déduire l'existence d'une base $\mathbf{e}$ de E telle que $\mathrm{Mat}(u\,;\mathbf{e})$ soit la matrice diagonale par blocs

$$\begin{pmatrix} C_\pi & & \\ & \ddots & \\ & & C_\pi \end{pmatrix}.$$

c. Montrer que deux endomorphismes $u, u' \in \mathscr{L}(E)$ tels que

$$\mu_u = \mu_{u'} = \pi$$

sont semblables. Le résultat subsiste-t-il si π n'est pas irréductible ?

E22. Un étudiant un peu paresseux n'a que trois activités journalières : manger, dormir, et chercher ses exos de maths. Toutes les cinq minutes, il peut soit changer d'activité, soit continuer celle qu'il était en train de faire.

Quand il dort, il a neuf chances sur dix de ne pas se réveiller dans les cinq minutes suivantes. Quand il se réveille, il y a une chance sur deux qu'il aille manger et une chance sur deux qu'il parte faire ses exercices. Le repas ne dure que cinq minutes, après il fait autre chose. Après avoir mangé, il y a trois chances sur dix qu'il parte faire ses exercices de maths, mais surtout sept chances sur dix qu'il retourne dormir. Réfléchir cinq minutes à ses exos est fatigant ; il y a huit chances sur dix qu'il retourne dormir au bout de cinq minutes. Sinon, il continue

en oubliant qu'il est déjà un peu fatigué[5]. On note D_n, M_n, E_n la probabilité que l'étudiant dorme, mange, fasse ses exos au bout de n périodes de cinq minutes. On note $X_n = \begin{pmatrix} D_n \\ M_n \\ E_n \end{pmatrix}$.

a. Montrer que $X_{n+1} = AX_n$ pour tout $n \geqslant 0$, avec

$$A = \begin{pmatrix} \frac{9}{10} & \frac{7}{10} & \frac{8}{10} \\ \frac{1}{20} & 0 & 0 \\ \frac{1}{2} & \frac{3}{10} & \frac{2}{10} \end{pmatrix}.$$

En déduire que $X_n = A^n \begin{pmatrix} D_0 \\ M_0 \\ E_0 \end{pmatrix}$.

On donne $\chi_A = (X-1)\left(X^2 - \frac{1}{10}X + \frac{1}{200}\right)$.

b. La matrice A est-elle diagonalisable sur $\mathbb{R}$? sur $\mathbb{C}$? Montrer qu'il existe un unique vecteur propre V_1 associé à la valeur propre 1 dont la somme des coordonnées vaut 1, et le calculer.

c. Diagonaliser A sur $\mathbb{C}$.

d. Lorsqu'un temps assez long s'est écoulé, donner une probabilité approchée que l'étudiant dorme. Cette probabilité dépend elle de X_0 ?

E23. Quelle est la décomposition de Dunford de

$$M = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} ?$$

E24.

a. Soit $E = \begin{pmatrix} 1 & -1 & 0 & 1 \\ 0 & 2 & -1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & -1 & 2 \end{pmatrix}$.

Montrer que E est diagonalisable et admet 1 et 2 comme valeurs propres.

b. Soit $A = \begin{pmatrix} 0 & -2 & 1 & 2 \\ -1 & 0 & 0 & 2 \\ -1 & -1 & 2 & 1 \\ -1 & -2 & 0 & 4 \end{pmatrix}$, et soit $F = A - E$.

Calculer F^2 et $EF - FE$.

c. Quelle est la décomposition de Dunford de A ?

5. Toute ressemblance avec des étudiants existants ou ayant existé est évidemment purement...

E25. Trouver la décomposition de Dunford de la matrice

$$M = \begin{pmatrix} 0 & 0 & 0 & -4 \\ 1 & 0 & 0 & 4 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{pmatrix} \in \mathrm{M}_4(\mathbb{C}).$$

E26. Soit K un corps, soit $A \in \mathrm{M}_n(K)$ et soit $B = \begin{pmatrix} 4A & A \\ -2A & A \end{pmatrix}$.

a. Diagonaliser $\begin{pmatrix} 4 & 1 \\ -2 & 1 \end{pmatrix}$.

b. Montrer que B est semblable à $\begin{pmatrix} 2A & 0 \\ 0 & 3A \end{pmatrix}$.

c. Trouver une condition nécessaire et suffisante sur A pour que B soit diagonalisable.

d. Donner la décomposition de Dunford de B en fonction de celle de A.

E27. Soit E un K-espace vectoriel de dimension $n \geqslant 1$, et soit $u \in \mathscr{L}(E)$. On note ℓ_u et r_u les applications

$$\begin{array}{rccc} \ell_u : & \mathscr{L}(E) & \longrightarrow & \mathscr{L}(E) \\ & f & \longmapsto & u \circ f \end{array} \quad \text{et} \quad \begin{array}{rccc} r_u : & \mathscr{L}(E) & \longrightarrow & \mathscr{L}(E) \\ & f & \longmapsto & f \circ u. \end{array}$$

On définit enfin l'application φ_u par

$$\begin{array}{rccc} \varphi_u : & \mathscr{L}(E) & \longrightarrow & \mathscr{L}(E) \\ & f & \longmapsto & u \circ f - f \circ u. \end{array}$$

a. Montrer que u, ℓ_u et r_u ont mêmes polynômes minimaux.

b. Montrer que si u est nilpotent/trigonalisable/diagonalisable, les endomorphismes ℓ_u, r_u et φ_u sont nilpotents/trigonalisables/diagonalisables.

Indication. On pourra remarquer que ℓ_u et r_u commutent.

c. On suppose que u est diagonalisable. Soit $(v_1, \ldots, v_n)$ une base de E formée de vecteurs propres de u. Pour tous $i, j \in [\![1, n]\!]$, soit e_{ij} l'unique endomorphisme de E tel que

$$e_{ij}(v_k) = \delta_{jk} v_i, \quad \text{pour tout } k \in [\![1, n]\!].$$

Montrer que la famille $(e_{ij})_{i,j \in [\![1,n]\!]}$ est une base de $\mathscr{L}(E)$, formée de vecteurs propres de φ_u.

d. On suppose que χ_u est scindé. Montrer que ℓ_u, r_u et φ_u ont des polynômes caractéristiques scindés (on ne demande pas de les calculer). Donner la décomposition de Dunford de ℓ_u, r_u et φ_u en fonction de celle de u.

E28. Soit K un corps de caractéristique différente de 2.

a. (i) Montrer qu'il existe une suite de polynômes $(P_m)_{m\geqslant 0}$ de $K[X]$ vérifiant les conditions suivantes pour tout $m \geqslant 0$:

▷ $\deg(P_m) = m$;

▷ $P_0 = 1$ et $P_{m+1} \equiv P_m \mod X^{m+1}$;

▷ $P_m^2 \equiv 1 + X \mod X^{m+1}$.

On pourra écrire $P_m^2 = 1 + X + Q_m X^{m+1}$, avec $Q_m \in K[X]$ et remarquer que $P_m(0) = 1$ pour tout $m \geqslant 0$.

(ii) Soit $\lambda \in K^\times$. On suppose qu'il existe $\mu \in K^\times$ tel que $\lambda = \mu^2$. Montrer qu'il existe une suite de polynômes $(f_{m,\lambda})_{m\geqslant 0}$ de $K[X]$ tel que $f_{m,\lambda}^2 \equiv X \mod (X-\lambda)^{m+1}$, pour tout $m \geqslant 0$.

b. Soit $r \geqslant 1$ un entier, et soit $N \in \mathrm{M}_r(K)$ une matrice nilpotente. On suppose qu'il existe $\mu \in K^\times$ tel que $\lambda = \mu^2$.

Calculer $\left(f_{r-1,\lambda}(\lambda \mathrm{I}_r + N)\right)^2$.

c. Soit $B \in \mathrm{GL}_n(\mathbb{C})$. Sans utiliser les résultats sur l'exponentielle de matrices, montrer qu'il existe $P \in \mathbb{C}[X]$ telle que $P(B)^2 = B$.

Indication. On pourra commencer par le cas d'un bloc de Jordan, puis utiliser le théorème chinois pour le cas général.

d. Soit $A = J_{n,0} \in \mathrm{M}_n(K)$, avec $n \geqslant 2$. Montrer qu'il n'existe pas de matrice $M \in \mathrm{M}_n(K)$ telle que $A = M^2$.

Indication. Si M existe, montrer que $A^{n-1} = 0$.

e. Soit $N \in \mathrm{M}_n(K)$ une matrice nilpotente.

Montrer qu'il existe $M \in \mathrm{M}_n(K)$ telle que $M^2 = N$ si, et seulement si, toutes les cellules de Jordan de N, sauf éventuellement une cellule de taille 1, peuvent être regroupées en paires de la forme $(J_{m,0}, J_{m,0})$ ou $(J_{m,0}, J_{m+1,0}), m \geqslant 1$.

f. Donner une condition nécessaire et suffisante pour que $A \in \mathrm{M}_n(\mathbb{C})$ ait une racine carrée.

E29. Soit $A \in \mathrm{M}_n(\mathbb{C})$.

a. Calculer la décomposition de Dunford de e^A en fonction de celle de A.

Indication. Montrer que si $N \in \mathrm{M}_n(\mathbb{C})$ est nilpotente, $e^N - \mathrm{I}_n$ est nilpotente.

b. Montrer que A est diagonalisable si, et seulement si, e^A est diagonalisable.

E30. Soit $K = \mathbb{R}$ ou $\mathbb{C}$. On dit que $A \in \mathrm{M}_n(K)$ est *unipotente* si $A - \mathrm{I}_n$ est nilpotente.

a. Justifier qu'une matrice unipotente est inversible.

b. Montrer que l'exponentielle induit une bijection entre l'ensemble des matrices nilpotentes et l'ensemble des matrices unipotentes.

E31. Soit u un endomorphisme de $\mathbb{R}^3$ tel que $\mu_u = (X-1)(X-2)^2$. Donner la forme de Jordan de u.

E32. Soient u, $u' \in \mathscr{L}(E)$ deux endomorphismes de polynômes caractéristiques scindés. Montrer que u et u' sont semblables si, et seulement si, on a

$$\dim_K \left(\mathrm{Ker}\big((u - \alpha \mathrm{Id}_E)^k\big)\right) = \dim_K \left(\mathrm{Ker}\big((u' - \alpha \mathrm{Id}_E)^k\big)\right)$$

pour tout $\alpha \in K$ et tout $k \geqslant 0$.

E33. Trouver une forme de Jordan et une base de Jordan pour les matrices suivantes :

$$\begin{pmatrix} 3 & 2 & -2 \\ -1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 4 & 2 & -2 \\ 1 & 5 & -1 \\ 1 & 3 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 & 2 \\ -3 & 0 & 4 & 0 \\ 0 & 1 & 0 & 3 \\ -1 & 0 & 1 & 0 \end{pmatrix}$$

$$\begin{pmatrix} -1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 2 & 2 \\ 0 & 1 & -1 & 1 \end{pmatrix}, \quad \begin{pmatrix} 3 & -1 & 1 & 7 \\ 9 & -3 & -7 & -1 \\ 0 & 0 & 4 & -8 \\ 0 & 0 & 2 & -4 \end{pmatrix}$$

$$\begin{pmatrix} -2 & -1 & 1 & 2 \\ 1 & -4 & 1 & 2 \\ 0 & 0 & -5 & 4 \\ 0 & 0 & -1 & -1 \end{pmatrix}, \quad \begin{pmatrix} 0 & -2 & 1 & 2 \\ -1 & 0 & 0 & 2 \\ -1 & -1 & 2 & 1 \\ -1 & -2 & 0 & 4 \end{pmatrix}.$$

Les exercices suivants utilisent la proposition ci-après, que le lecteur en L3 pourra utiliser librement (cf. proposition XXVI-1.6).

Proposition. Soit K un corps, et soit $f \in K[X]$ non constant. Alors, il existe un corps L contenant K tel que f soit scindé dans $L[X]$.

E34. Soit K un corps et soit $n \geqslant 1$ un entier. Soit $M \in \mathrm{M}_n(K)$.

a. Justifier l'existence d'un corps L contenant K sur lequel M est trigonalisable.

b. Montrer que l'on a alors

$$\mathrm{tr}(M^k) = \sum_{\lambda \in \mathrm{Sp}_L(M)} m_\lambda \lambda^k, \quad \text{pour tout } k \geqslant 1.$$

c. En déduire que si M est nilpotente, on a $\mathrm{tr}(M^k) = 0$ pour tout $k \geqslant 1$. Retrouver ce résultat de manière plus directe, sans utiliser l'existence de L et la question **b.**

d. On suppose que $\mathrm{car}(K) > n$. Soit $M \in \mathrm{M}_n(K)$ telle que $\mathrm{tr}(M^k) = 0$ pour tout $k \geqslant 1$. Montrer que M est nilpotente.

Indication. Soit L un corps contenant K sur lequel χ_M est scindé. Supposer que M possède au moins une valeur propre non nulle dans L. Soient $\lambda_1, \ldots, \lambda_s \in L^\times$ les valeurs propres non nulles de M, deux à deux distinctes, et soient $m_1, \ldots, m_s$ leurs multiplicités respectives. Utiliser **b** pour montrer que le vecteur $\begin{pmatrix} m_1 \\ \vdots \\ m_s \end{pmatrix}$ est solution d'un système linéaire dont le déterminant est non nul.

e. On suppose que $\mathrm{car}(K) \leqslant n$. Le résultat de **d** est-il encore vrai ?

E35. Soit K un corps, et soit $A \in \mathrm{M}_n(K)$. On se propose d'étudier l'équation

$$AM - MA = A, \quad M \in \mathrm{M}_n(K). \qquad (*)$$

On note $C(A)$ l'ensemble des matrices de $\mathrm{M}_n(K)$ qui commutent avec A.

a. Montrer que si l'équation $(*)$ a une solution $M_0 \in \mathrm{M}_n(K)$, l'ensemble des solutions est $M_0 + C(A)$.

Le but de ce qui suit est d'étudier l'existence d'une solution à l'équation $(*)$. On commence par étudier le cas où A est nilpotente.

b. On suppose que $A = J_{n,0}$.

(i) Montrer que $C(A) = K[A]$.

Indication. Soit $(\varepsilon_1, \ldots, \varepsilon_n)$ la base canonique de K^n. Si $M \in C(A)$, et si $M\varepsilon_n = a_1\varepsilon_1 + \cdots + a_n\varepsilon_n$, montrer que

$$M = a_1 A^{n-1} + \cdots + a_{n-1}A + a_n \mathrm{I}_n.$$

Pour cela, on pourra démontrer et utiliser les relations

$$A^{n-k}\varepsilon_n = \varepsilon_k, \quad \text{pour tout } k \in [\![1, n]\!].$$

(ii) Soit T (resp. T^0) le sous-espace de $\mathrm{M}_n(K)$ dont les éléments sont les matrices triangulaires supérieures (resp. les matrices triangulaires supérieures à diagonale nulle).

Vérifier que si $M \in T$, alors $AM - MA \in T^0$. Montrer alors que l'application linéaire

$$\begin{aligned} T &\longrightarrow T^0 \\ M &\longmapsto AM - MA \end{aligned}$$

est surjective, et en déduire que $(*)$ possède au moins une solution.

Indication. Utiliser le théorème du rang.

c. On suppose que A est nilpotente. Montrer soigneusement que l'équation $(*)$ possède au moins une solution.

On s'intéresse maintenant à la réciproque. On se propose de démontrer de deux manières différentes le résultat suivant : si $\mathrm{car}(K) > n$, et si l'équation $(*)$ possède au moins une solution, alors A est nilpotente.

d. Montrer que pour tout entier $k \geqslant 1$, on a $\mathrm{tr}(A^k) = 0$. Conclure en utilisant l'exercice précédent.

e. On se propose ici d'utiliser une autre méthode.

(i) Montrer que pour tout $P \in K[X]$, on a

$$P(A)M - MP(A) = AP'(A).$$

Indication. Traiter d'abord le cas où $P = X^k$.

(ii) En déduire que $\mu_A \mid X\mu'_A$, puis que $d\mu_A = X\mu'_A$, où $d = \deg(\mu_A)$ (on fera bien attention de souligner à quel endroit l'hypothèse sur K est utilisée).

(iii) En déduire que $\mu_A = X^d$, et conclure.

f. On suppose que $K = \mathbb{F}_2$, que $n = 2$, et que $A = \begin{pmatrix} \overline{1} & \overline{1} \\ \overline{0} & \overline{1} \end{pmatrix}$.

Résoudre l'équation $(*)$. La matrice A est-elle nilpotente ?

E36. Soit $M = (a_{ij}) \in \mathrm{M}_n(K)$. Pour toute partie I de $[\![1, n]\!]$ (éventuellement vide), on note $|I|$ le nombre d'éléments de I. On note aussi M_I la matrice extraite $M_I = (a_{k\ell})_{k,\ell \in I}$.

On se propose ici de démontrer l'identité

$$\chi_M = \sum_{I \subset [\![1,n]\!]} (-1)^{|I|} \det(M_I) X^{n-|I|},$$

avec la convention $\det(M_\emptyset) = 1$.

a. Justifier que l'on peut supposer que χ_M est scindé.

b. Montrer l'égalité pour une cellule de Jordan.

c. Montrer que si M est diagonale par blocs, et si l'égalité est vraie pour chacun des blocs diagonaux, alors elle est vraie pour M.

d. Montrer l'égalité dans le cas général.

E37. Soit K un corps de caractéristique nulle, et soit $N \in \mathrm{M}_n(K)$.

a. Montrer que N et $2N$ sont semblables si, et seulement si, N est nilpotente.

Indication. Pour « $\Longleftarrow$ », on pourra remarquer qu'une matrice nilpotente possède toujours une forme de Jordan. Pour « $\Longrightarrow$ », considérer un corps L contenant K sur lequel le polynôme χ_N est scindé, puis montrer que si $\lambda \in L$ est valeur propre de N, alors 2λ l'est aussi.

b. Soit $N = \begin{pmatrix} \overline{0} & \overline{1} \\ -\overline{1} & \overline{0} \end{pmatrix} \in \mathrm{M}_2(\mathbb{F}_3)$. Montrer que N et $2N$ sont semblables. Commenter.

E38. On souhaite montrer que toute matrice carrée complexe est semblable à une matrice symétrique.

Si $m \geqslant 1$, on note $Q_m \in \mathrm{M}_m(\mathbb{C})$ la matrice possédant des 1 sur l'antidiagonale, et des zéros partout ailleurs.

a. Pour tout $\lambda \in \mathbb{C}$, et tout $n \geqslant 1$, montrer que $P^{-1} J_{n,\lambda} P$ est symétrique, où $P = \begin{pmatrix} Q_m & -iQ_m \\ \mathrm{I}_m & i\mathrm{I}_m \end{pmatrix}$ si $n=2m$, et $P = \begin{pmatrix} Q_m & 0 & -iQ_m \\ 0 & \sqrt{2} & 0 \\ \mathrm{I}_m & 0 & i\mathrm{I}_m \end{pmatrix}$ si $n=2m+1$.

b. Conclure.

Chapitre XXXI

Décomposition de Frobenius

Dans le chapitre précédent, nous avons montré comment la décomposition de Jordan permettait de résoudre le problème de classification des endomorphismes à similitude près. Néanmoins, le fait de devoir calculer les valeurs propres d'un endomorphisme et leurs multiplicités est plutôt problématique, puisque cela revient à savoir factoriser complètement le polynôme caractéristique.

Dans ce chapitre, nous allons démontrer l'existence d'une décomposition qui caractérise les endomorphismes d'un espace vectoriel de dimension finie sur un corps quelconque à similitude près, et qui peut se calculer de manière totalement algorithmique.

1. Sous-espace stable engendré par un vecteur

Nous allons nous intéresser dans cette section à une catégorie particulière de sous-espaces stables. Dans toute la suite, K sera un corps quelconque, et E sera un K-espace vectoriel non nul de dimension finie.

1.1. Définition. Soit $u \in \mathscr{L}(E)$. Si $v \in E$, le *sous-espace stable engendré par* v est le plus petit sous-espace de E stable par u contenant v. Autrement dit, c'est l'intersection de tous les sous-espaces [1] de E stables par u contenant v.

C'est un sous-espace de E, stable par u et contenant v. On le note E_v.

1. Remarquons qu'il existe au moins un tel sous-espace, à savoir E lui-même.

1.2. Lemme. *Soit $u \in \mathscr{L}(E)$. Pour tout $v \in E$, on a*

$$E_v = \{P(u)(v) \mid P \in K[X]\}.$$

De plus, une base de E_v est donnée par

$$\big(v,\ u(v),\ \ldots,\ u^{d-1}(v)\big),$$

où $d \geqslant 0$ est le plus petit entier tel que la famille $\big(v,\ u(v),\ \ldots,\ u^d(v)\big)$ soit liée.

Par ailleurs, pour tout $P \in K[X]$, on a

$$P(u_{E_v}) = 0 \iff P(u)(v) = 0.$$

Enfin, si $a_0, \ldots, a_{d-1} \in K$ vérifient $u^d(v) = -a_0 v - a_1 u(v) - \cdots - a_{d-1}u^{d-1}(v)$, alors

$$\mu_{u_{E_v}} = X^d + a_{d-1}X^{d-1} + \cdots + a_1 X + a_0.$$

En particulier, $\deg(\mu_{u_{E_v}}) = \dim_K(E_v)$.

Démonstration. Gardons les notations de l'énoncé.

Il est facile de vérifier que $\{P(u)(v) \mid P \in K[X]\}$ est un sous-espace de E stable par u contenant v. Par définition, on a donc

$$E_v \subset \{P(u)(v) \mid P \in K[X]\}.$$

Par ailleurs, puisque E_v est stable par u, il est stable par tout polynôme en u. Comme $v \in E_v$, on en déduit

$$\{P(u)(v) \mid P \in K[X]\} \subset E_v,$$

d'où l'égalité recherchée.

Puisque E est de dimension finie, il existe un plus petit entier $d \geqslant 0$ tel que la famille

$$\big(v,\ u(v),\ \ldots,\ u^d(v)\big)$$

soit liée. Nous allons vérifier que $\mathbf{e} = \big(v,\ u(v),\ \ldots,\ u^{d-1}(v)\big)$ est une base de E_v. Par choix de d, $\mathbf{e}$ est libre. Comme la famille $\big(v,\ u(v), \ldots, u^d(v)\big)$ est liée, il existe donc des scalaires $\lambda_0, \ldots, \lambda_d \in K$ non tous nuls tels que

$$\lambda_0 v + \cdots + \lambda_{d-1}u^{d-1}(v) + \lambda_d u^d(v) = 0.$$

Si λ_d était nul, alors $\lambda_0, \ldots, \lambda_{d-1}$ seraient également nuls puisque la famille $\big(v, u(v),\ \ldots,\ u^{d-1}(v)\big)$ est libre. En divisant par λ_d, on obtient donc l'existence de scalaires $a_0, \ldots, a_{d-1} \in K$ tels que

$$u^d(v) = -a_0 v - a_1 u(v) - \cdots - a_{d-1}u^{d-1}(v).$$

Posons $P_0 = X^d + a_{d-1}X^{d-1} + \cdots + a_1 X + a_0 \in K[X]$. Pour tout $m \geqslant 0$, il existe Q_m, $R_m \in K[X]$, $\deg(R_m) \leqslant d-1$ tels que $X^m = Q_m P_0 + R_m$.

Par définition de P_0, on a $P_0(u)(v) = 0$. On a donc

$$u^m(v) = Q_m(u)\big(P_0(u)(v)\big) + R_m(u)(v) = R_m(u)(v)$$

pour tout $m \geqslant 0$. Comme $\deg(R_m) \leqslant d-1$, $u^m(v) = R_m(v)$ est donc une combinaison linéaire des vecteurs $v,\ u(v),\ \ldots,\ u^{d-1}(v)$.

Cela entraîne facilement que $\mathbf{e}$ est génératrice, donc une base.

Soit $P \in K[X]$. Si $P(u_{E_v}) = 0$, alors

$$P(u_{E_v})(v) = P(u)_{E_v}(v) = P(u)(v) = 0.$$

Inversement, si $P\big(u(v)\big) = 0$, alors pour tout $m \in [\![0, d-1]\!]$, on a

$$P(u)\big(u^m(v)\big) = u^m\big(P(u)(v)\big) = u^m(0) = 0.$$

Ainsi, $P(u)$ s'annule sur une base de E_v, donc sur E_v tout entier. Autrement dit,

$$P(u_{E_v}) = P(u)_{E_v} = 0.$$

Enfin, avec les notations précédentes, $\mu_{u_{E_v}} \mid P_0$. En effet, on a $P_0(u)(v) = 0$ par définition de P_0, et l'on conclut en utilisant l'équivalence précédente et les propriétés du polynôme minimal.

Supposons que $d' = \deg(\mu_{u_{E_v}}) < d$. La division euclidienne de X^{d-1} par $\mu_{u_{E_v}}$ montre comme précédemment que $u^{d-1}(v)$ est une combinaison linéaire de $v,\ u(v),\ \ldots,\ u^{d'-1}(v)$. En particulier, $\big(v,\ u(v),\ \ldots,\ u^{d-1}(v)\big)$ est une famille liée, d'où une contradiction avec le choix de d. Ainsi, $\mu_{u_{E_v}}$ et P_0 ont même degré. Comme $\mu_{u_{E_v}} \mid P_0$ et que ces deux polynômes sont unitaires, on en déduit que $\mu_{u_{E_v}} = P_0$. Cela achève la démonstration. □

Puisque E_v est un sous-espace stable par u, on a $\mu_{u_{E_v}} \mid \mu_u$. Le résultat suivant montre qu'il existe un vecteur v pour lequel on a en fait égalité.

1.3. Proposition. *Soit $u \in \mathscr{L}(E)$. Alors, il existe un vecteur $v \in E \setminus \{0\}$ tel que $\mu_{u_{E_v}} = \mu_u$.*

Démonstration. Supposons dans un premier temps que $\mu_u = \pi^m$, où $m \geqslant 1$ et $\pi \in K[X]$ est irréductible unitaire.

Remarquons que $\operatorname{Ker}\big(\pi^{m-1}(u)\big) \neq E$. Sinon, π^{m-1} serait un polynôme annulateur de u dont le degré est strictement inférieur à celui de μ_u, d'où une contradiction. Soit alors $v \in E \setminus \operatorname{Ker}\big(\pi^{m-1}(u)\big)$. Montrons que v convient. Si $\mu_{u_{E_v}}$ était un diviseur strict de μ_u, on aurait

$$\mu_{u_{E_v}} \mid \pi^{m-1},$$

et, par suite, on aurait $\pi^{m-1}(u)(x) = 0$ pour tout vecteur $x \in E_v$, c'est-à-dire $E_v \subset \operatorname{Ker}\big(\pi^{m-1}(u)\big)$.

En particulier, on obtiendrait $v \in \mathrm{Ker}\big(\pi^{m-1}(u)\big)$, ce qui est exclu par choix de v. On a donc bien $\mu_{u_{E_v}} = \mu_u$ dans ce cas.

Revenons au cas général, et écrivons

$$\mu_u = \pi_1^{m_1} \cdots \pi_r^{m_r},$$

où $s \geqslant 1$, $m_i \geqslant 1$ et $\pi_1, \ldots, \pi_r$ sont irréductibles unitaires, deux à deux distincts. Pour $i \in [\![1, r]\!]$, posons $F_i = \mathrm{Ker}\big(\pi_i^{m_i}(u)\big)$. Par le lemme de décomposition des noyaux, on a

$$E = F_1 \oplus \cdots \oplus F_r.$$

Par le corollaire XXX-1.21, on a $\mu_{u_{F_i}} = \pi_i^{m_i}$, pour tout $i \in [\![1, r]\!]$. D'après le point précédent, il existe $v_i \in F_i$ tel que $\mu_{u_{E_{v_i}}} = \mu_{u_{F_i}} = \pi_i^{m_i}$. Posons alors $v = v_1 + \cdots + v_r$.

Soit $P \in K[X]$. Par le lemme 1.2, on a

$$\mu_{u_{E_v}} \mid P \iff P(u_{E_v}) = 0 \iff P(u)(v) = 0.$$

Or, $P(u)(v) = P(u)(v_1) + \cdots + P(u)(v_r) = 0$. Le sous-espace F_i étant stable par u, on a $P(u)(v_i) \in F_i$. Comme $E = F_1 \oplus \cdots \oplus F_r$, on en déduit que

$$P(u_{E_v}) = 0 \iff P(u)(v_i) = 0_{E_{v_i}}, \quad \text{pour tout } i \in [\![1, r]\!].$$

Par le lemme 1.2, on en déduit alors la suite d'équivalences

$$\begin{aligned} P(u_{E_v}) = 0 &\iff P(u_{E_{v_i}}) = 0, \quad \text{pour tout } i \in [\![1, r]\!] \\ &\iff \pi_i^{m_i} \mid P, \quad \text{pour tout } i \in [\![1, r]\!] \\ &\iff \mu_u \mid P. \end{aligned}$$

Finalement, pour tout $P \in K[X]$, on a $\mu_{u_{E_v}} \mid P$ si, et seulement si, $\mu_u \mid P$, et donc $\mu_{u_{E_v}} = \mu_u$. Cela achève la démonstration. □

1.4. Définition. Soit $u \in \mathscr{L}(E)$. Un vecteur $v \in E \setminus \{0\}$ tel que $\mu_{u_{E_v}} = \mu_u$ est appelé un vecteur *u-maximal.*

2. Endomorphismes cycliques

Nous allons maintenant introduire la notion cruciale de ce chapitre.

2.1. Définition. Soit $u \in \mathscr{L}(E)$. On dit que u est un endomorphisme *cyclique* s'il existe $v \in E$ tel que la famille

$$\big(v,\ u(v),\ \ldots,\ u^{n-1}(v)\big)$$

soit une base de E. Un tel vecteur v est dit *u-cyclique.*

2.2. Remarque. Si u est un endomorphisme cyclique, et si $v \in E$ est un vecteur u-cyclique, alors il existe des scalaires $a_0, \dots, a_{n-1} \in K$ uniques tels que

$$u^n(v) = -a_0 v - a_1 u(v) - \dots - a_{n-1} u^{n-1}(v).$$

La matrice représentative de u dans la base $\big(v,\ u(v),\ \dots,\ u^{n-1}(v)\big)$ est alors

$$\begin{pmatrix} 0 & 0 & & 0 & -a_0 \\ 1 & 0 & & & -a_1 \\ 0 & 1 & & & -a_2 \\ & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}.$$

Rappelons que, dans ce cas, on a $\chi_u = X^n + a_{n-1}X^{n-1} + \dots + a_1 X + a_0$ par l'exemple XXX-1.13.

Inversement, supposons qu'il existe une base $\mathbf{e} = (e_1, \dots, e_n)$ de E telle que

$$\mathrm{Mat}(u\,;\mathbf{e}) = \begin{pmatrix} 0 & 0 & & 0 & -a_0 \\ 1 & 0 & & & -a_1 \\ 0 & 1 & & & -a_2 \\ & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}.$$

Alors, u est cyclique, et e_1 est u-cyclique. En effet, par définition d'une matrice représentative, on a

$$u(e_i) = e_{i+1}, \qquad \text{pour tout } i \in [\![1, n-1]\!].$$

On en déduit par récurrence que $e_j = u^j(e_1)$, pour tout $j \in [\![0, n-1]\!]$, d'où le résultat.

Cela invite à donner un nom aux matrices de cette forme particulière.

2.3. Définition. Soit $P = X^n + a_{n-1}X^{n-1} + \dots + a_1 X + a_0 \in K[X]$ un polynôme unitaire. La *matrice compagnon* de P est la matrice

$$C_P = \begin{pmatrix} 0 & 0 & & 0 & -a_0 \\ 1 & 0 & & & -a_1 \\ 0 & 1 & & & -a_2 \\ & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix} \in \mathrm{M}_n(K).$$

Remarquons que si $n = 1$, et $P = X + a_0$, on a $C_P = (-a_0)$.

2.4. Remarque. D'après la remarque 2.2, $u \in \mathscr{L}(E)$ est donc cyclique si, et seulement si, il existe une base $\mathbf{e} = (e_1, \ldots, e_n)$ telle que la matrice représentative $\mathrm{Mat}(u\,;\mathbf{e})$ soit une matrice compagnon. Dans ce cas, le vecteur e_1 est un vecteur u-cyclique.

Nous allons maintenant caractériser les endomorphismes cycliques.

2.5. Proposition. *Soit E un K-espace vectoriel de dimension $n \geqslant 1$, et soit $u \in \mathscr{L}(E)$. Les propriétés suivantes sont équivalentes :*

(1) *l'endomorphisme u est cyclique ;*

(2) *il existe une base $\mathbf{e}$ telle que $\mathrm{Mat}(u\,;\mathbf{e})$ soit une matrice compagnon ;*

(3) *il existe $v \in E \setminus \{0\}$ tel que $E = E_v$;*

(4) *on a $\deg(\mu_u) = n$;*

(5) *on a $\chi_u = \mu_u$.*

Démonstration

L'équivalence (1) $\iff$ (2) a été démontrée dans la remarque 2.2, et l'équivalence (4) $\iff$ (5) est aisée : comme χ_u et μ_u sont tous deux unitaires et $\mu_u \mid \chi_u$, ces deux polynômes sont égaux si, et seulement si, ils ont même degré. Comme $\deg(\chi_u) = n$, on en déduit le résultat.

Si maintenant $E = E_v$, $v \neq 0$ puisque E est non nul, et E admet une base de la forme voulue d'après le lemme 1.2. Ainsi, u est cyclique. Inversement, si u est cyclique, et si v est un vecteur cyclique, alors $v \neq 0$ (puisque c'est un élément d'une base de E), et E_v contient une base de E par hypothèse. Ainsi, $E = E_v$, d'où l'équivalence (1) $\iff$ (3).

Il reste donc à établir l'équivalence (3) $\iff$ (4), par exemple. Si $E = E_v$, on a évidemment $\mu_u = \mu_{u_{E_v}}$. Le lemme 1.2 montre alors que l'on a les égalités

$$n = \dim_K(E_v) = \deg(\mu_u).$$

Réciproquement, supposons que l'on ait $\deg(\mu_u) = n$. D'après la proposition 1.3, il existe un vecteur $v \in E \setminus \{0\}$ tel que $\mu_{u_{E_v}} = \mu_u$, puis, par le lemme 1.2, E_v est alors un sous-espace de E de dimension n, et donc $E = E_v$, d'où le résultat. □

2.6. Exemple. Dès que $\dim_K(E) \geqslant 2$, il existe des endomorphismes non cycliques, comme l'exemple du morphisme Id_E le montre immédiatement.

En revanche, on peut construire facilement des sous-espaces F stables par un endomorphisme $u \in \mathscr{L}(E)$ tels que $u_F \in \mathscr{L}(F)$ soit cyclique. Par le lemme 1.2, il suffit par exemple de prendre $F = E_v$, où $v \in E \setminus \{0\}$.

3. Décomposition de Frobenius : approche classique

Nous allons maintenant étudier l'existence d'une décomposition de E en somme directe de sous-espaces stables sur lesquels la restriction de u est cyclique. Comme nous l'avons remarqué dans la section précédente, pour tout $v \in E\backslash\{0\}$, E_v est un exemple de tel sous-espace. L'idée est de procéder par récurrence. Pour cela, il faudrait s'assurer que tout sous-espace stable possède un supplémentaire stable (ou tout du moins, que cela soit le cas pour les sous-espaces stables engendrés par un vecteur v).

Hélas, ce n'est pas toujours vrai, comme le montre l'exemple qui suit.

3.1. Exemple. Soit $E = K^2$, soit $(\varepsilon_1, \varepsilon_2)$ sa base canonique, et soit u l'endomorphisme de E dont la matrice représentative dans la base canonique est

$$M = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Alors, E_{ε_1} n'admet pas de supplémentaire stable par u. Sinon, un tel supplémentaire serait une droite stable, donc engendré par un vecteur propre de u, et M serait diagonalisable, ce qui n'est pas le cas.

Les endomorphismes $u \in \mathscr{L}(E)$ tels que tout sous-espace stable par u admette un supplémentaire stable sont étudiés dans l'exercice XXX-20.

Fort heureusement, nous avons le résultat crucial suivant.

3.2. Lemme. *Soit $u \in \mathscr{L}(E)$, et soit $v \in E$ un vecteur u-maximal*[2]*. Alors, le sous-espace E_v admet un supplémentaire stable par u.*

Démonstration. Soit v un vecteur u-maximal, et soit $F = E_v$.

Par le lemme 1.2, il existe $d \geqslant 0$ tel que

$$\big(v,\ u(v),\ \ldots,\ u^{d-1}(v)\big)$$

soit une base de E_v. Toujours par ce même lemme, puisque v est u-maximal, on a

$$\dim_K(F) = \deg(\mu_u) = d.$$

Complétons cette famille libre en une base $\mathbf{e} = (e_0, \ldots, e_{n-1})$ de E, de sorte que

$$e_k = u^k(v), \qquad \text{pour tout } k \in [\![0, d-1]\!].$$

2. Vecteur dont l'existence est assurée par la proposition 1.3.

Posons $G = \bigcap_{k=0}^{d-1} \mathrm{Ker}\big(e_{d-1}^* \circ u^k\big)$. Nous allons montrer que G est le supplémentaire recherché.

Tout d'abord, G est stable par u. En effet, soit $x \in G$. On a donc

$$e_{d-1}^*\big(u^k(x)\big) = 0, \quad \text{pour tout } k \in [\![0, d-1]\!].$$

La division euclidienne de X^{k+1} par μ_u montre que u^{k+1} est combinaison linéaire de Id_E, u, $\ldots$, u^{d-1}, et cela pour tout $k \in [\![0, d-1]\!]$. Ainsi, par linéarité de e_{d-1}^*, pour tout $k \in [\![0, d-1]\!]$, $e_{d-1}^* \circ u^{k+1}$ est combinaison linéaire de e_{d-1}^*, $\ldots$, $e_{d-1}^* \circ u^{d-1}$. On en déduit que, pour tout $k \in [\![0, d-1]\!]$, on a

$$(e_{d-1}^* \circ u^k)\big(u(x)\big) = (e_{d-1}^* \circ u^{k+1})(x) = 0,$$

et donc $u(x) \in G$.

Montrons maintenant que $F \cap G = \{0\}$. Soit $x \in F \cap G$. Supposons $x \neq 0$. On peut alors écrire

$$x = \lambda_0 e_0 + \cdots + \lambda_k e_k,$$

avec $k \in [\![0, d-1]\!]$ et $\lambda_k \neq 0$. Or, pour tout $i \in [\![0, k]\!]$, on a

$$(e_{d-1}^* \circ u^{d-1-k})(e_i) = (e_{d-1}^* \circ u^{d-1-k+i})(v) = e_{d-1}^*(e_{d-1-(k-i)}) = \delta_{ik}.$$

Comme $x \in G$ et que $d-1-k \in [\![0, d-1]\!]$, on obtient alors

$$(e_{d-1}^* \circ u^{d-1-k})(x) = 0 = \lambda_k,$$

d'où une contradiction. On a donc bien $F \cap G = \{0\}$.

Il reste à voir que $\dim_K(F) + \dim_K(G) = \dim_K(E)$. Montrons pour cela que e_{d-1}^*, $e_{d-1}^* \circ u$, $\ldots$, $e_{d-1}^* \circ u^{d-1}$ sont linéairement indépendantes.

Supposons le contraire. Alors, il existe $\lambda_0, \ldots, \lambda_{d-1} \in K$ non tous nuls tels que

$$\lambda_0 e_{d-1}^* + \lambda_1 e_{d-1}^* \circ u + \cdots + \lambda_{d-1} e_{d-1}^* \circ u^{d-1} = 0.$$

Soit $k \in [\![0, d-1]\!]$ le plus grand entier tel que $\lambda_k \neq 0$. On a ainsi

$$\lambda_0 e_{d-1}^* + \lambda_1 e_{d-1}^* \circ u + \cdots + \lambda_k e_{d-1}^* \circ u^k = 0,$$

avec $\lambda_k \in K^\times$. En appliquant e_{d-1-k} à cette égalité, et en remarquant que l'on a $u^i(e_{d-1-k}) = e_{d-1-(k-i)}$, pour tout $i \in [\![0, k]\!]$, on obtient $\lambda_k = 0$, d'où une contradiction. Par le lemme III-5.10, on a

$$\dim_K(G) = n - d = \dim_K(E) - \dim_K(F),$$

ce qui achève la démonstration. □

Nous pouvons à présent énoncer le théorème principal de ce chapitre.

3.3. Théorème (Décomposition de Frobenius). *Soit E un K-espace vectoriel de dimension $n \geqslant 1$, et soit $u \in \mathscr{L}(E)$.*

Alors, il existe un entier $r \geqslant 1$, des sous-espaces $F_1, \dots, F_r$ et des polynômes unitaires non constants $\chi_{u,1}, \dots, \chi_{u,r} \in K[X]$ vérifiant les conditions suivantes :

(1) *on a $E = F_1 \oplus \cdots \oplus F_r$;*

(2) *pour tout $i \in [\![1, r]\!]$, F_i est stable par u, et u_{F_i} est cyclique, de polynôme minimal $\chi_{u,i}$;*

(3) *on a $\chi_{u,1} \mid \chi_{u,2} \mid \cdots \mid \chi_{u,r}$.*

Une telle décomposition est appelée une décomposition de Frobenius *de u. L'entier r et les polynômes $\chi_{u,1}, \dots, \chi_{u,r}$ sont uniques, et appelés les* invariants de similitude *de u. On a les relations*

$$\chi_{u,r} = \mu_u \quad \text{et} \quad \chi_{u,1} \cdots \chi_{u,r} = \chi_u.$$

De plus, il existe une base $\mathbf{e}$ de E telle que

$$\mathrm{Mat}(u\,;\mathbf{e}) = \begin{pmatrix} C_{\chi_{u,1}} & & \\ & \ddots & \\ & & C_{\chi_{u,r}} \end{pmatrix}.$$

Enfin, deux endomorphismes de E sont semblables si, et seulement si, ils ont mêmes invariants de similitude.

Démonstration. Commençons par l'existence d'une telle décomposition. On va procéder par récurrence. Plus précisément, pour tout $n \geqslant 1$, soit (H_n) la proposition suivante :

(H_n) Pour tout K-espace vectoriel E non nul de dimension $\leqslant n$, tout endomorphisme de E admet une décomposition de Frobenius.

Supposons tout d'abord que $n = 1$. Alors, $\dim_K(E) = 1$ puisque E est non nul, et tout endomorphisme de E est cyclique. On pose alors $r = 1$, $F_1 = E$ et $\chi_{u,1} = \mu_u$. Ainsi, (H_1) est vraie.

Soit $n \geqslant 1$. Supposons que (H_n) soit vraie, et montrons que (H_{n+1}) est vraie. Si $\dim_K(E) \leqslant n$, on applique (H_n), et l'on a fini. On peut donc supposer que $\dim_K(E) = n + 1$. Soit $v \in E$ un vecteur u-maximal. En particulier, on a $v \neq 0$. Alors, E_v est stable par u, et $u_{E_v} \in \mathscr{L}(E_v)$ est cyclique. Par le lemme précédent, E_v admet un supplémentaire G stable par u. Si G est nul, on a $E = E_v$, et l'on a fini. Si G est non nul, comme E_v est non nul, on a alors $\dim_K(G) \leqslant n$ et l'on peut appliquer l'hypothèse de récurrence pour trouver une décomposition de Frobenius de u_G, disons

$$G = F_1 \oplus \cdots \oplus F_{r-1}.$$

Alors, pour tout $i \in [\![1, r-1]\!]$, $u_{F_i} = (u_G)_{F_i}$ est cyclique.

Posons $F_r = E_v$, et pour tout $i \in [\![1, r]\!]$, notons $\chi_{u,i}$ le polynôme minimal de u_{F_i}. Pour voir que $E = F_1 \oplus \cdots \oplus F_r$ est une décomposition de Frobenius de u, il reste à vérifier que $\chi_{u,1} | \chi_{u,2} | \cdots | \chi_{u,r}$. Si $i \in [\![1, r-1]\!]$, cela provient du fait que $\chi_{u,i}$ est le polynôme minimal de $u_{F_i} = (u_G)_{F_i}$, et que

$$G = F_1 \oplus \cdots \oplus F_{r-1}$$

est une décomposition de Frobenius de u_G. Enfin, le vecteur v étant u-maximal, on a donc $\chi_{u,r} = \mu_u$. Comme μ_u annule $u_{F_{r-1}}$, on a $\mu_{u_{F_{r-1}}} \mid \mu_u$, c'est-à-dire $\chi_{u,r-1} \mid \chi_{u,r}$. Cela achève la récurrence, et la démonstration de l'existence d'une décomposition de Frobenius.

Soit $E = F_1 \oplus \cdots \oplus F_r$ une décomposition de Frobenius de u. Pour tout entier $i \in [\![1, r]\!]$, choisissons une base $\mathbf{e}_i$ de F_i telle que $\mathrm{Mat}(u_{F_i}; \mathbf{e}_i)$ soit une matrice compagnon, associé à un polynôme P_i. Ce polynôme P_i est alors le polynôme caractéristique de u_{F_i}, c'est-à-dire également son polynôme minimal par la proposition 2.5. On a donc

$$\mathrm{Mat}(u_{F_i}; \mathbf{e}_i) = C_{\chi_{u,i}}.$$

En recollant ces bases, on obtient une base $\mathbf{e}$ de E telle que

$$\mathrm{Mat}(u\,; \mathbf{e}) = \begin{pmatrix} C_{\chi_{u,1}} & & \\ & \ddots & \\ & & C_{\chi_{u,r}} \end{pmatrix}.$$

On en déduit alors immédiatement la relation

$$\chi_u = \chi_{u,1} \cdots \chi_{u,r}.$$

En tenant compte du fait que $\chi_{u,1} \mid \cdots \mid \chi_{u,r}$, le lemme XXX-1.20 montre alors que l'on a

$$\mu_u = \mathrm{ppcm}(\chi_{u,1}, \ldots, \chi_{u,r}) = \chi_{u,r}.$$

Passons à la partie délicate de la démonstration, à savoir l'unicité de la décomposition.

Soient $E = F_1 \oplus \cdots \oplus F_r = G_1 \oplus \cdots \oplus G_s$ deux décompositions de Frobenius de u. Pour tout $k \in [\![1, r]\!]$, on note P_k le polynôme minimal de u_{F_k}, et pour tout $\ell \in [\![1, s]\!]$, on note Q_ℓ le polynôme minimal de u_{G_j}.

Sans perte de généralité, on peut supposer que $s \geqslant r$.

Soit $j \in [\![1, r]\!]$ le plus grand entier tel que

$$P_{r-i+1} = Q_{s-i+1}, \quad \text{pour tout } i \in [\![1, j]\!].$$

Notons qu'un tel entier existe, puisque $P_r = \mu_u = Q_s$. De, plus, par définition de j, on a $P_{r-j} \neq Q_{s-j}$, sauf si $j = r$, auquel cas cette condition est vide. Dans la suite, afin de ne pas avoir à considérer ce cas particulier,

nous poserons $P_0 = 1$. Le lecteur vérifiera alors que les arguments présentés ci-dessous sont encore valables.

Comme les sous-espaces $F_1, \dots, F_r$ sont stables par u, on a

$$P_{r-j}(u)(F_i) \subset F_i, \quad \text{pour tout } i \in [\![1, r]\!].$$

En particulier, $P_{r-j}(u)(F_1), \dots, P_{r-j}(u)(F_r)$ sont aussi en somme directe. Remarquons aussi que

$$P_{r-j}(u)(F_1) = \cdots = P_{r-j}(u)(F_{r-j}) = 0,$$

puisque pour tout $k \in [\![1, r-j]\!]$, $P_k = \mu_{u_{F_k}}$ divise P_{r-j}. En appliquant alors l'endomorphisme $P_{r-j}(u)$ à $E = F_1 \oplus \cdots \oplus F_r$, on obtient

$$P_{r-j}(u)(E) = P_{r-j}(u)(F_{r-j+1}) \oplus \cdots \oplus P_{r-j}(u)(F_r).$$

De même, on a

$$P_{r-j}(u)(E) = P_{r-j}(u)(G_1) \oplus \cdots \oplus P_{r-j}(u)(G_s).$$

Par hypothèse, pour tout $i \in [\![1, j]\!]$, on a $P_{r-i+1} = Q_{s-i+1}$. Ainsi, il existe une base $\mathbf{f}_{r-i+1}$ de F_{r-i+1} et une base $\mathbf{g}_{s-i+1}$ de G_{s-i+1} telles que

$$\mathrm{Mat}(u_{F_{r-i+1}}; \mathbf{f}_{r-i+1}) = C_{P_{r-i+1}} = C_{Q_{s-i+1}} = \mathrm{Mat}(u_{G_{s-i+1}}; \mathbf{g}_{s-i+1}).$$

Mais alors,

$$\mathrm{Mat}\big(P_{r-j}(u_{F_{r-i+1}}); \mathbf{f}_{r-i+1}\big) = P_{r-j}(C_{P_{r-i+1}}) = P_{r-j}(C_{Q_{s-i+1}}),$$

ce qui se récrit

$$\mathrm{Mat}\big(P_{r-j}(u_{F_{r-i+1}}); \mathbf{f}_{r-i}\big) = \mathrm{Mat}\big(P_{r-j}(u_{G_{s-i+1}}); \mathbf{g}_{s-i+1}\big).$$

Par conséquent, $P_{r-j}(u_{F_{r-i+1}})$ et $P_{r-j}(u_{G_{s-i+1}})$ ont même rang, d'où

$$\dim_K\big(P_{r-j}(u)(F_{r-i+1})\big) = \dim_K\big(P_{r-j}(u)(G_{s-i+1})\big), \quad \text{pour tout } i \in [\![1, j]\!].$$

On en déduit alors

$$\dim_K\big(P_{r-j}(u)(G_1)\big) + \cdots + \dim_K\big(P_{r-j}(u)(G_{s-j})\big) = 0.$$

En particulier, $P_{r-j}(u)(G_{s-j}) = 0$, ou encore $P_{r-j}(u_{G_{s-j}}) = 0$. On en déduit donc que $Q_{s-j} \mid P_{r-j}$. En appliquant cette fois le polynôme Q_{s-j} aux deux décompositions, on obtient $P_{r-j} \mid Q_{s-j}$, d'où $P_{r-j} = Q_{s-j}$, et une contradiction si $j < r$. Ainsi, on a $j = r$, et donc

$$P_{r-i+1} = Q_{s-i+1}, \quad \text{pour tout } i \in [\![1, r]\!].$$

Par conséquent, $\dim_K(F_{r-i+1}) = \dim_K(G_{s-i+1})$, pour tout $i \in [\![1, r]\!]$.

Si l'on avait $s > r$, on aurait alors

$$\dim_K(G_1) = \cdots = \dim_K(G_{s-r}) = 0,$$

soit $G_1 = \cdots = G_{s-r} = \{0\}$, ce qui est impossible. Ainsi, $r = s$, et l'on a bien l'unicité.

Il reste à montrer la dernière partie. Soient u, $u' \in \mathscr{L}(E)$ deux endomorphismes semblables. Il existe donc $w \in \mathrm{GL}(E)$ tel que $u = w \circ u' \circ w^{-1}$. Soient $E = F_1 \oplus \cdots \oplus F_r$ et $E = G_1 \oplus \cdots \oplus G_s$ des décompositions de Frobenius de u et u' respectivement.

L'égalité $u = wu'w^{-1}$ montre que $w(G_1), \ldots, w(G_s)$ sont stables par u. De plus, w étant un isomorphisme, on a aisément

$$E = w(G_1) \oplus \cdots \oplus w(G_s).$$

Enfin, $u_{w(G_j)}$ est cyclique, de même polynôme minimal (ou caractéristique) que u'_{G_j}. En effet, si $\big(v_j,\ u'(v_j),\ \ldots,\ u'^{m_j-1}(v_j)\big)$ est une base de G_j, alors la famille

$$\big(w(v_j),\ w\big(u'(v_j)\big),\ \ldots,\ w\big(u'^{m_j-1}(v_j)\big)\big)$$

est une base de $w(G_j)$. Mais $wu' = uw$, et l'on en déduit donc que

$$\big(w(v_j),\ u\big(w(v_j)\big),\ \ldots,\ u^{m_j-1}\big(w(v_j)\big)\big)$$

est une base de $w(G_j)$. L'égalité des polynômes minimaux provient du fait que w est bijective.

Ainsi, $E = w(G_1) \oplus \cdots \oplus w(G_s)$ est une décomposition de Frobenius de u, dont les facteurs invariants sont les facteurs invariants de u'. Par unicité des facteurs invariants, on en déduit que u et u' ont mêmes facteurs invariants. Réciproquement, si u et u' ont mêmes facteurs invariants, par un point précédent, il existe deux bases $\mathbf{e} = (e_1, \ldots, e_n)$ et $\mathbf{e}' = (e'_1, \ldots, e'_n)$ de E telles que $\mathrm{Mat}(u\,;\mathbf{e}) = \mathrm{Mat}(u'\,;\mathbf{e}')$. L'unique endomorphisme $w \in \mathscr{L}(E)$ vérifiant

$$w(e_i) = e'_i, \qquad \text{pour tout } i \in [\![1, n]\!]$$

est un automorphisme de E (puisqu'il envoie une base de E sur une base de E), et l'on a $w \circ u' \circ w^{-1} = u$. Par conséquent, u et u' sont semblables. Cela achève la démonstration. □

Bien que la démonstration donne un moyen pratique d'obtenir une décomposition de Frobenius, la méthode mise en jeu n'est pas très aisée et nécessite a priori de savoir calculer de manière efficace un polynôme minimal[3]. C'est pourquoi nous nous proposons d'établir dans la suite un algorithme de calcul, basé sur les résultats du chapitre XX. La théorie cachée derrière est celle des modules de type fini sur un anneau principal.

3. Cela afin de calculer un vecteur u-maximal.

Néanmoins, nous ne souhaitons pas développer cela dans cet ouvrage, et nous renvoyons le lecteur à [9] pour un exposé plus conceptuel. Cette référence explique également comment passer d'une décomposition de Frobenius à une décomposition de Jordan, et vice versa, ce que nous ne ferons pas ici, vu la longueur déjà démesurée de ce livre. Néanmoins, le lecteur pourra se référer aux exercices de ce chapitre à ce sujet.

4. Décomposition de Frobenius : calcul pratique

Le but de ce qui suit est de présenter un algorithme de calcul des invariants de similitude d'un endomorphisme, ainsi qu'une décomposition de Frobenius. Cet algorithme va utiliser l'algorithme de réduction des matrices à coefficients dans une anneau euclidien présenté dans le chapitre XX (cf. section XX-2).

On ne suppose pas ici que les résultats du théorème 3.3 soient établis, puisque nous les redémontrerons au passage.

La proposition suivante fait le lien entre invariants de similitude et équivalence de matrices de $\mathrm{M}_n(K[X])$.

4.1. Proposition. *Soit E un K-espace vectoriel de dimension $n \geqslant 1$, et soit $u \in \mathscr{L}(E)$. On note M la matrice représentative de u dans une base fixée de E.*

On suppose que u admet une décomposition de Frobenius

$$E = F_1 \oplus \cdots \oplus F_r.$$

Si $\chi_{u,i} \in K[X]$ est le polynôme minimal de u_{F_i}, alors la matrice $X\mathrm{I}_n - M$ est $K[X]$-équivalente à

$$\begin{pmatrix} \mathrm{I}_{n-r} & & \\ & \chi_{u,1} & \\ & & \ddots \\ & & & \chi_{u,r} \end{pmatrix}.$$

De plus, $n - r + 1 = \min\{j \in [\![1, n]\!] \mid \mu_j(X\mathrm{I}_n - M) \neq 1\}$, et

$$\chi_{i,f} = \frac{\mu_{n-r+i}(X\mathrm{I}_n - M)}{\mu_{n-r+i-1}(X\mathrm{I}_n - M)}, \quad \text{pour tout } i \in [\![1, r]\!],$$

où $\mu_j(X\mathrm{I}_n - M)$ est le pgcd *unitaire des j-mineurs de $X\mathrm{I}_n - M$ (avec la convention $\mu_0(X\mathrm{I}_n - M) = 1$).*

En particulier, l'entier r et les polynômes $\chi_{u,1}, \ldots, \chi_{u,r}$ sont uniques.

Démonstration. Comme dans la démonstration du théorème 3.3, on établit l'existence d'une base $\mathbf{e}$ telle que

$$\mathrm{Mat}(u\,;\mathbf{e}) = \begin{pmatrix} C_{\chi_{u,1}} & & \\ & \ddots & \\ & & C_{\chi_{u,r}} \end{pmatrix}.$$

Notons C cette dernière matrice.

Il existe donc $P \in \mathrm{GL}_n(K)$ telle que $P^{-1}MP = C$. On a ainsi

$$P^{-1}(X\mathrm{I}_n - M)P = X\mathrm{I}_n - C = \begin{pmatrix} X\mathrm{I}_{d_1} - C_{\chi_{u,1}} & & \\ & \ddots & \\ & & X\mathrm{I}_{d_r} - C_{\chi_{u,r}} \end{pmatrix},$$

où $d_i = \deg(\chi_{u,i})$.

Supposons avoir montré le résultat lorsque M est une matrice compagnon. Il existe donc des matrices inversibles S_i, $T_i \in \mathrm{GL}_{d_i}(K[X])$ telles que

$$S_i(X\mathrm{I}_{d_i} - C_{\chi_{u,i}})T_i = \begin{pmatrix} \mathrm{I}_{d_i-1} & \\ & P \end{pmatrix}, \quad \text{pour tout } i \in [\![1, r]\!].$$

Si l'on pose $S = \begin{pmatrix} S_1 & & \\ & \ddots & \\ & & S_r \end{pmatrix}$ et $T = \begin{pmatrix} T_1 & & \\ & \ddots & \\ & & T_r \end{pmatrix}$, alors S et T sont deux éléments de $\mathrm{GL}_n(K[X])$, et l'on a

$$S(X\mathrm{I}_n - C)T = \begin{pmatrix} \mathrm{I}_{d_1-1} & & & & \\ & \chi_{u,1} & & & \\ & & \ddots & & \\ & & & \mathrm{I}_{d_r-1} & \\ & & & & \chi_{u,r} \end{pmatrix}.$$

En multipliant à gauche par un produit de matrices de transpositions convenables (ce qui revient à réordonner les éléments diagonaux), on obtient une matrice $Q \in \mathrm{GL}_n(K)$ telle que $QS(X\mathrm{I}_n - C)T$ soit de la forme voulue. Comme $\mathrm{GL}_n(K) \subset \mathrm{GL}_n(K[X])$, on en déduit aisément que $X\mathrm{I}_n - M$ est alors $K[X]$-équivalente à

$$\begin{pmatrix} \mathrm{I}_{n-r} & & & \\ & \chi_{u,1} & & \\ & & \ddots & \\ & & & \chi_{u,r} \end{pmatrix}.$$

Il reste à montrer le résultat pour une matrice compagnon.

Soit $P = X^n + a_{n-1}X^{n-1} + \cdots + a_1X + a_0 \in K[X]$, et soit C_P la matrice compagnon associée. On a

$$X\mathrm{I}_n - C_P = \begin{pmatrix} X & 0 & & 0 & a_0 \\ -1 & X & & & a_1 \\ 0 & -1 & & & a_2 \\ & & \ddots & & \vdots \\ 0 & 0 & \cdots & -1 & X + a_{n-1} \end{pmatrix} \in \mathrm{M}_n(K[X]).$$

L'opération $L_1 \longleftarrow L_1 + XL_2 + \cdots + X^{n-1}L_n$, qui est une succession d'opérations élémentaires, montre que l'on a

$$X\mathrm{I}_n - C_P \sim \begin{pmatrix} 0 & 0 & & 0 & P \\ -1 & X & & & a_1 \\ 0 & -1 & & & a_2 \\ & & \ddots & & \vdots \\ 0 & 0 & \cdots & -1 & X + a_{n-1} \end{pmatrix}.$$

En effectuant les opérations $C_j \longleftarrow C_j + XC_{j-1}$, pour $j \in [\![1, n-1]\!]$, on obtient

$$X\mathrm{I}_n - C_P \sim \begin{pmatrix} 0 & 0 & & 0 & P \\ -1 & 0 & & & a_1 \\ 0 & -1 & & & a_2 \\ & & \ddots & & \vdots \\ 0 & 0 & \cdots & -1 & a_{n-1} \end{pmatrix}.$$

Finalement, l'opération $C_n \longleftarrow a_1C_1 + a_2C_2 + \cdots + a_{n-1}C_{n-1} + C_n$ montre que

$$X\mathrm{I}_n - C \sim \begin{pmatrix} 0 & 0 & & 0 & P \\ -1 & 0 & & & 0 \\ 0 & -1 & & & 0 \\ & & \ddots & & \vdots \\ 0 & 0 & \cdots & -1 & 0 \end{pmatrix}.$$

En échangeant successivement les colonnes de cette dernière matrice, on obtient que $X\mathrm{I}_n - C_P$ est $K[X]$-équivalente à $\begin{pmatrix} -\mathrm{I}_{n-1} & \\ & P \end{pmatrix}$.

En alternant échange de colonnes et de lignes et multiplication par la matrice de dilatation $D^{(n)}(-1)$, on remplace tous les -1 de la diagonale par des 1, et l'on a le résultat désiré.

La dernière partie découle du théorème XX-2.3, appliqué à l'anneau euclidien $K[X]$. □

Maintenant que nous savons comment calculer les invariants de similitude, nous allons nous intéresser à l'obtention pratique d'une décomposition de Frobenius et d'une base de Frobenius associée, tout en redémontrant leur existence. On commence par deux résultats techniques.

4.2. Lemme. *Soit $u \in \mathscr{L}(E)$ et soit $\mathbf{e} = (e_1, \dots, e_n)$ une base de E. Soit $M = \mathrm{Mat}(u\,;\mathbf{e}) \in \mathrm{M}_n(K)$, et soient $C_1, \dots, C_n$ les colonnes de la matrice $X\mathrm{I}_n - M \in \mathrm{M}_n(K[X])$.*

Alors, l'application K-linéaire

$$\begin{aligned} \varphi : K[X]^n &\longrightarrow E \\ \begin{pmatrix} P_1 \\ \vdots \\ P_n \end{pmatrix} &\longmapsto P_1(u)(e_1) + \cdots + P_n(u)(e_n) \end{aligned}$$

est surjective, et $\mathrm{Ker}(\varphi)$ est l'ensemble des combinaisons $K[X]$-linéaires[4] *des colonnes de $X\mathrm{I}_n - M$, c'est-à-dire*

$$\mathrm{Ker}(\varphi) = \{Q_1C_1 + \cdots + Q_nC_n \mid Q_1, \dots, Q_n \in K[X]\}.$$

Démonstration. On vérifie aisément que φ est bien K-linéaire. De plus, il est clair que e_i est dans l'image de φ pour tout $i \in [\![1, n]\!]$. On en déduit immédiatement la surjectivité de φ, puisque son image contient une base de E. Vérifions que $C_1, \dots, C_n \in \mathrm{Ker}(\varphi)$. En effet, on a

$$\varphi(C_j) = \sum_{i=1}^{n} (\delta_{ij}X - a_{ij})(u)(e_i) = \sum_{i=1}^{n} \big(\delta_{ij}u(e_i) - a_{ij}e_i\big) = u(e_j) - \sum_{i=1}^{n} a_{ij}e_i.$$

Mais, par définition de M, on a

$$u(e_j) = \sum_{i=1}^{n} a_{ij}e_i.$$

Ainsi, on a $C_1, \dots, C_n \in \mathrm{Ker}(\varphi)$. Remarquons maintenant que pour tout polynôme $P \in K[X]$ et tout $v = \begin{pmatrix} P_1 \\ \vdots \\ P_n \end{pmatrix} \in K[X]^n$, on a

$$\varphi(Pv) = P(u)\big(\varphi(v)\big),$$

comme le montre un simple calcul. En particulier, si $v \in \mathrm{Ker}(\varphi)$, il en est de même de Pv pour tout $P \in K[X]$. On en déduit alors aisément l'inclusion

$$\{Q_1C_1 + \cdots + Q_nC_n \mid Q_1, \dots, Q_n \in K[X]\} \subset \mathrm{Ker}(\varphi).$$

4. Dans le langage des $K[X]$-modules, $\mathrm{Ker}(\varphi)$ est donc le sous-module de $K[X]^n$ engendré par les colonnes de $X\mathrm{I}_n - M$.

Soit maintenant $v = \begin{pmatrix} P_1 \\ \vdots \\ P_n \end{pmatrix} \in \mathrm{Ker}(\varphi)$. Nous allons montrer par récurrence sur $d = \max_i \big(\deg(P_i)\big)$ que v est combinaison linéaire de $C_1, \dots, C_n$. Si tous les P_i sont nuls, il n'y a rien à faire. Si $d = 0$, c'est-à-dire si $P_i = a_i \in K$, pour tout $i \in [\![1, n]\!]$, alors l'égalité

$$P_1(u)(e_1) + \cdots + P_n(u)(e_n) = a_1 e_1 + \cdots + a_n e_n = 0$$

implique que $P_1 = \cdots = P_n = 0$, car $\mathbf{e}$ est une K-base de E. Ainsi, $v = 0$ et le résultat est encore trivial.

Supposons maintenant que l'on ait prouvé la propriété jusqu'au rang $d-1$, et prouvons-la au rang d. Effectuons à cet effet la division euclidienne de P_i par $X - a_{ii}$. On a donc

$$P_i = (X - a_{ii})Q_i + c_i, \quad \text{où } c_i \in K.$$

Par construction, on a $\deg(Q_i) \leqslant d-1$, pour tout entier $i \in [\![1, n]\!]$. Or, on a $v - \sum_{j=1}^{n} Q_j C_j \in \mathrm{Ker}(\varphi)$, d'après le point précédent, et de plus la i-ème coordonnée de $v - \sum_{j=1}^{n} Q_j C_j$ est $P_i - \sum_{j=1}^{n} (\delta_{ij} X - a_{ij}) Q_j$, par choix des P_i et par définition de $C_1, \dots, C_n$. Par ailleurs,

$$P_i - \sum_{j=1}^{n} (\delta_{ij} X - a_{ij}) Q_j = P_i - (X - a_{ii}) Q_i - \sum_{j \neq i} a_{ij} Q_j = c_i - \sum_{j \neq i} a_{ij} Q_j.$$

On a alors

$$\deg\Big(P_i - \sum_{j=1}^{n} (\delta_{ij} X - a_{ij}) Q_j\Big) \leqslant d-1, \quad \text{pour tout } i \in [\![1, n]\!].$$

Par hypothèse de récurrence, $v - \sum_{j=1}^{n} Q_j C_j$ est donc une combinaison linéaire des C_j, et par suite v l'est également. □

4.3. Corollaire. *Soit* $\mathbf{e}$ *une base de* E, *et soit* $M = \mathrm{Mat}(u\,; \mathbf{e})$. *Soient* $U, V \in \mathrm{GL}_n(K[X])$ *telles que*

$$U(X\mathrm{I}_n - M)V = \begin{pmatrix} \mathrm{I}_{n-r} & & \\ & P_1 & \\ & & \ddots \\ & & & P_r \end{pmatrix},$$

où les polynômes $P_1, \dots, P_r$ *sont unitaires non constants et* $P_1 \mid \cdots \mid P_r$. *Soient* $Y_1, \dots, Y_r \in K[X]^n$ *les* r *dernières colonnes de* U^{-1}. *Écrivons*

$$Y_j = \begin{pmatrix} R_{1j} \\ \vdots \\ R_{nj} \end{pmatrix}, \quad \textit{pour tout } j \in [\![1, r]\!],$$

et posons

$$v_j = \sum_{i=1}^{n} R_{ij}(u)(e_i), \quad \textit{pour tout } j \in [\![1,r]\!].$$

Pour tous $Q_1, \dots, Q_r \in K[X]$, *on a*

$$Q_1(u)(v_1) + \cdots + Q_r(u)(v_r) = 0 \implies P_j \mid Q_j, \quad \textit{pour tout } j \in [\![1,r]\!].$$

Démonstration. Soient des polynômes $Q_1, \dots, Q_r \in K[X]$ vérifiant la relation précédente. Par définition de v_j, on a alors

$$\sum_{j=1}^{r}\sum_{i=1}^{n} Q_j(u)\big(R_{ij}(u)(e_i)\big) = \sum_{i=1}^{n}\Big(\sum_{j=1}^{r}(Q_j R_{ij})(u)\Big)(e_i) = 0.$$

Remarquons que, si $Z_1, \dots, Z_{n-r}, Y_1, \dots, Y_r \in K[X]^n$ désignent les colonnes de U^{-1}, cette égalité se retraduit par

$$\varphi(Q_1Y_1 + \cdots + Q_rY_r) = 0,$$

où φ désigne l'application définie dans le lemme 4.2. Par ce même lemme, on en déduit $Q_1Y_1 + \cdots + Q_rY_r$ est une combinaison $K[X]$-linéaire de colonnes de $X\mathrm{I}_n - M$. Remarquons que les colonnes de la matrice

$$U^{-1}\begin{pmatrix} \mathrm{I}_{n-r} & & & \\ & P_1 & & \\ & & \ddots & \\ & & & P_r \end{pmatrix} V^{-1}$$

sont $Z_1, \dots, Z_{n-r}, P_1Y_1, \dots, P_rY_r$. La relation

$$X\mathrm{I}_n - M = U^{-1}\begin{pmatrix} \mathrm{I}_{n-r} & & & \\ & P_1 & & \\ & & \ddots & \\ & & & P_r \end{pmatrix} V^{-1}$$

montre alors que les colonnes de la matrice $X\mathrm{I}_n - M$ sont des combinaisons $K[X]$-linéaires de $Z_1, \dots, Z_{n-r}, P_1Y_1, \dots, P_rY_r$.

Il existe donc des polynômes $S_1, \dots, S_{n-r}, T_1, \dots, T_r \in K[X]$ tels que

$$Q_1Y_1 + \cdots + Q_rY_r = S_1Z_1 + \cdots + S_{n-r}Z_{n-r} + T_1P_1Y_1 + \cdots + T_rP_rY_r.$$

Comme U est inversible dans $\mathrm{M}_n(K[X])$, donc dans $\mathrm{M}_n\big(K(X)\big)$, ses colonnes sont $K(X)$-linéairement indépendantes. En particulier, on obtient

$$Q_i = T_iP_i, \quad \text{pour tout } i \in [\![1,r]\!],$$

d'où le résultat. □

On peut maintenant décrire la méthode explicite souhaitée.

4.4. Proposition. *Soit E un K-espace vectoriel de dimension $n \geqslant 1$, et soit $u \in \mathscr{L}(E)$. Soit $\mathbf{e} = (e_1, \dots, e_n)$ une base de E, et soit $M = \mathrm{Mat}(u\,;\mathbf{e})$. Soient U, $V \in \mathrm{GL}_n(K[X])$ deux matrices telles que*

$$U(X\mathrm{I}_n - M)V = \begin{pmatrix} \mathrm{I}_{n-r} & & & \\ & P_1 & & \\ & & \ddots & \\ & & & P_r \end{pmatrix},$$

où $P_1, \dots, P_r \in K[X]$ sont unitaires non constants et vérifient les relations de divisibilité $P_1 \mid P_2 \mid \cdots \mid P_r$.

Soient $Y_1, \dots, Y_r \in K[X]^n$ les r dernières colonnes de U^{-1}. Écrivons

$$Y_j = \begin{pmatrix} R_{1j} \\ \vdots \\ R_{nj} \end{pmatrix}, \quad \text{pour tout } j \in [\![1, r]\!],$$

et posons

$$v_j = \sum_{i=1}^{n} R_{ij}(u)(e_i), \quad \text{pour tout } j \in [\![1, r]\!].$$

Si $d_j = \deg(P_j)$, et si $F_j = \mathrm{Vect}_K\big(v_j, u(v_j), \dots, u^{d_j-1}(v_j)\big)$, alors

$$E = F_1 \oplus \cdots \oplus F_r$$

est une décomposition de Frobenius de u, et le polynôme minimal de u_{F_j} est égal à P_j.
En particulier, $P_1, \dots, P_r$ sont les invariants de similitude de u.

Démonstration. Gardons les notations de l'énoncé, et montrons que la famille de vecteurs

$$\big(v_1, u(v_1), \ \dots, \ u^{d_1-1}(v_1), \ \dots, \ v_r, u(v_r), \ \dots, \ u^{d_r-1}(v_r)\big)$$

est libre. Les combinaisons K-linéaires de ces vecteurs sont exactement les éléments de E de la forme

$$Q_1(u)(v_1) + \cdots + Q_r(u)(v_r),$$

où $\deg(Q_i) < d_i$, pour tout $i \in [\![1, r]\!]$.

D'après le corollaire 4.3, une telle combinaison linéaire est nulle si, et seulement si, $P_i \mid Q_i$, pour tout entier i dans $[\![1, r]\!]$. Comme $\deg(P_i) = d_i$, on obtient $Q_1 = \cdots = Q_r = 0$, ce qui revient à dire les coefficients de la combinaison linéaire considérée sont tous nuls.

La famille précédente est donc libre. Pour démontrer que c'est une base, il suffit de vérifier qu'elle est de cardinal n, c'est-à-dire la relation

$$\deg(P_1) + \cdots + \deg(P_r) = n.$$

La dernière partie du théorème XX-2.3, appliquée à $A = K[X]$ et à la

matrice $C = X\mathrm{I}_n - M$ montre que l'on a

$$P_1 \cdots P_r = \det(X\mathrm{I}_n - M) = \chi_M.$$

Ce polynôme étant de degré n, on obtient la relation demandée. Au vu de la définition de $F_1, \ldots, F_r$, il est alors clair que

$$E = F_1 \oplus \cdots \oplus F_r,$$

que chaque F_j est stable par u et que u_{F_j} est cyclique.

Calculons le polynôme minimal de u_{F_j}. D'après le lemme 1.2, pour tout polynôme $Q \in K[X]$, on a $Q(u_{F_j}) = 0$ si, et seulement si, $Q(u)(v_j) = 0$. Par le corollaire 4.3, cela est équivalent à $P_j \mid Q_j$. Ainsi, le polynôme minimal de u_{F_j} est P_j. Comme $P_1 \mid \cdots \mid P_r$, on a donc bien obtenu une décomposition de Frobenius de l'endomorphisme u. □

4.5. Corollaire. *Soit $u \in \mathscr{L}(E)$. Alors, on a*

$$\mu_u = \frac{\chi_M}{\mu_{n-1}(X\mathrm{I}_n - M)},$$

où $n = \dim_K(E)$, M est une matrice représentative de u dans une base fixée de E, et $\mu_{n-1}(X\mathrm{I}_n - M)$ est le pgcd unitaire des mineurs d'ordre $n-1$ de la matrice $X\mathrm{I}_n - M$.

Démonstration. Par le théorème 3.3, le dernier invariant de similitude est le polynôme minimal de u. Pour conclure, il suffit alors d'utiliser la proposition 4.1. □

4.6. Remarque. Ce dernier résultat donne donc une formule explicite pour le polynôme minimal de u. Elle est bien évidemment inutilisable en pratique, car très coûteuse en calculs[5]. Elle permet aussi de retrouver le théorème de Cayley-Hamilton : puisque $\chi_M = \chi_u$, l'énoncé donne en particulier que $\mu_u \mid \chi_u$, c'est-à-dire que χ_u annule u.

Tout cela doit paraître très miraculeux. Comme nous l'avons déjà expliqué, la théorie cachée derrière cette méthode explicite est celle des modules de type fini sur un anneau principal, théorie qui sort du cadre de cet ouvrage. Nous renvoyons le lecteur intéressé à [9] pour plus de détails.

Résumons ce qui précède : soit E un K-espace vectoriel de dimension $n \geqslant 1$, soit $u \in \mathscr{L}(E)$, et soit $M = \mathrm{Mat}(u\,;\mathbf{e})$, où $\mathbf{e}$ est une base de E fixée. L'algorithme de réduction des matrices à coefficients dans un anneau euclidien donne l'existence de deux matrices inversibles $U, V \in \mathrm{GL}_n(K[X])$ telles

5. Il faut calculer, non seulement le polynôme caractéristique, mais aussi n^2 déterminants de taille $n-1$, puis leur pgcd.

que

$$U(X\mathrm{I}_n - M)V = \begin{pmatrix} \mathrm{I}_{n-r} & & \\ & P_1 & \\ & & \ddots \\ & & & P_r \end{pmatrix},$$

où $P_1, \ldots, P_r \in K[X]$ sont unitaires non constants et vérifient les relations de divisibilité $P_1 \mid P_2 \mid \cdots \mid P_r$.

La matrice U (resp. V) est le produit de matrices de transvection, de transposition ou de dilatation correspondant aux opérations élémentaires effectuées sur les lignes (resp. les colonnes) de $X\mathrm{I}_n - M$.

(1) Si l'on veut seulement calculer les invariants de similitude de u, les résultats précédents montrent que les invariants de similitude de u sont obtenus en appliquant l'algorithme de réduction à la matrice $X\mathrm{I}_n - M$. Il n'est pas alors nécessaire de calculer les matrices U et V.

Les invariants de similitude de u sont alors

$$\chi_{u,1} = P_1, \ \ldots, \ \chi_{u,r} = P_r.$$

On peut même calculer les invariants de similitude en utilisant la formule

$$\chi_{i,f} = \frac{\mu_{n-r+i}(X\mathrm{I}_n - M)}{\mu_{n-r+i-1}(X\mathrm{I}_n - M)}, \quad \text{pour tout } i \in [\![1, r]\!],$$

où r est le plus petit entier tel que $\mu_{n-r+1}(X\mathrm{I}_n - M)$ est non constant, mais cela est très coûteux en calculs, sauf pour $n = 2$ ou 3.

(2) Supposons que l'on veuille calculer explicitement une base de Frobenius, et soient $Y_1, \ldots, Y_r \in K[X]^n$ les r dernières colonnes de U^{-1}. Écrivons

$$Y_j = \begin{pmatrix} R_{1j} \\ \vdots \\ R_{nj} \end{pmatrix}, \quad \text{pour tout } j \in [\![1, r]\!],$$

et posons

$$v_j = \sum_{i=1}^{n} R_{ij}(u)(e_i), \quad \text{pour tout } j \in [\![1, r]\!].$$

Si $d_j = \deg(P_j)$, alors

$$\left(v_1, \ u(v_1), \ \ldots, \ u^{d_1-1}(v_1), \ \ldots, \quad v_r, \ u(v_r), \ldots, \ u^{d_r-1}(v_r)\right)$$

est une base de Frobenius.

Il faut donc calculer U, ou plutôt U^{-1}. Si les opérations sur les lignes de $X\mathrm{I}_n - M$ correspondent successivement à la multiplication à gauche par des matrices $T_1, \ldots, T_s \in \mathrm{GL}_n(K[X])$, alors $U = T_s T_{s-1} \cdots T_1$, et donc

$$U^{-1} = T_1^{-1} \ldots T_s^{-1}.$$

Or, chaque matrice T_k est une matrice de transvection, de transposition ou de dilatation, dont l'inverse est très facile à calculer.

On a donc intérêt, d'une part à privilégier les opérations sur les colonnes plutôt que sur les lignes, et cela afin de minimiser les calculs, et d'autre part de calculer U^{-1} de proche en proche à chaque étape de l'algorithme.

4.7. Exemple. Soit $u \in \mathscr{L}(\mathbb{Q}^3)$ l'endomorphisme de $\mathbb{Q}^3$ dont la matrice représentative dans la base canonique est

$$M = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -4 & -1 & -1 \end{pmatrix}.$$

On a donc

$$X\mathrm{I}_3 - M = \begin{pmatrix} X-1 & 0 & 0 \\ 0 & X-1 & 0 \\ 4 & 1 & X+1 \end{pmatrix}.$$

En calculant les mineurs de cette matrice, il est facile de voir que l'on a

$$\mu_1(X\mathrm{I}_3-M)=1, \quad \mu_2(X\mathrm{I}_3-M)=X-1, \quad \mu_3(X\mathrm{I}_3-M)=(X-1)^2(X+1).$$

On a donc $\chi_{u,1} = X-1$ et $\chi_{u,2} = X^2-1$.

Puisque la matrice compagnon de $X-1$ est (1), et que la matrice compagnon de X^2-1 est $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, on en déduit que M est semblable à la matrice

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Nous allons maintenant trouver une matrice de passage explicite en utilisant l'algorithme. Bien sûr, vu la simplicité de l'exemple, on pourrait la trouver à la main, mais nous voulons surtout ici illustrer la méthode générale.

On veut calculer U^{-1} de proche en proche. On initialise par $U^{-1} = \mathrm{I}_3$. Au lieu d'indexer les matrices successivement obtenues, on gardera le même nom par abus de notation à chaque étape.

On calculera aussi directement les inverses des matrices utilisées. On a

$$X\mathrm{I}_3 - M \sim \begin{pmatrix} 0 & X-1 & 0 \\ X-1 & 0 & 0 \\ 1 & 4 & X+1 \end{pmatrix} \sim \begin{pmatrix} 1 & 4 & X+1 \\ X-1 & 0 & 0 \\ 0 & X-1 & 0 \end{pmatrix}$$

et donc

$$U^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Ensuite, on obtient

$$XI_3 - M \sim \begin{pmatrix} 1 & 0 & 0 \\ X-1 & 4-4X & 1-X^2 \\ 0 & X-1 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & 4-4X & 1-X^2 \\ 0 & X-1 & 0 \end{pmatrix}$$

et ainsi

$$U^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ X-1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ X-1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

On a ensuite

$$XI_3 - M \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & X-1 & 0 \\ 0 & 4-4X & 1-X^2 \end{pmatrix}.$$

On a donc

$$U^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ X-1 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ X-1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

On a alors

$$XI_3 - M \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & X-1 & 0 \\ 0 & 0 & 1-X^2 \end{pmatrix}.$$

On a donc

$$U^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ X-1 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -4 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ X-1 & -4 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Enfin, on a

$$XI_3 - M \sim \begin{pmatrix} 1 & 0 & 0 \\ 0 & X-1 & 0 \\ 0 & 0 & X^2-1 \end{pmatrix},$$

en multipliant la dernière colonne par -1, ce qui ne change pas U.

On a donc finalement

$$U^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ X-1 & -4 & 1 \\ 1 & 0 & 0 \end{pmatrix}.$$

Notons alors

$$v_1 = \begin{pmatrix} 1 \\ -4 \\ 0 \end{pmatrix} \quad \text{et} \quad v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}.$$

Posons

$$v_3 = f(v_2) = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}.$$

La proposition 4.4 nous assure alors que

$$f(v_1) = v_1, \quad f(v_2) = v_3 \quad \text{et} \quad f(v_3) = v_2,$$

ce que l'on peut vérifier à la main. La matrice de f dans la base (v_1, v_2, v_3) est donc

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}.$$

4.8. Remarque. Le théorème 3.3 et les méthodes explicites précédentes nous donnent un moyen extrêmement pratique de déterminer si deux endomorphismes u et u' de E sont semblables. Il suffit de calculer leurs invariants de similitudes. En outre, dans ce cas, on peut trouver explicitement un automorphisme $\varphi \in \mathrm{GL}(E)$ tel que $u' = \varphi \circ u \circ \varphi^{-1}$.

Fixons une base $\mathbf{e}$ de E, et soient M et M' les matrices représentatives de u et u' dans la base $\mathbf{e}$. Il suffit alors d'utiliser la méthode précédente pour trouver des bases de Frobenius $\mathbf{f}$ et $\mathbf{f}'$ associées à u et u'. Soit Q la matrice de passage de $\mathbf{f}$ à $\mathbf{e}$ et soit Q' la matrice de passage de $\mathbf{f}'$ à $\mathbf{e}$.

Puisque u et u' ont mêmes invariants de similitudes, leurs formes de Frobenius sont égales, et l'on a donc

$$\mathrm{Mat}(u\,;\mathbf{f}) = \mathrm{Mat}(u'\,;\mathbf{f}'),$$

soit encore

$$Q^{-1}MQ = Q'^{-1}M'Q'.$$

On prend alors pour φ l'unique automorphisme de E tel que

$$\mathrm{Mat}(\varphi\,;\mathbf{e}) = Q'Q^{-1}.$$

Tous les résultats précédents ont, bien sûr, leurs équivalents matriciels.

Si $M \in \mathrm{M}_n(K)$, on définit ses invariants de similitude comme étant les invariants de similitude de l'endomorphisme

$$\begin{aligned} u_M : K^n &\longrightarrow K^n \\ X &\longmapsto MX. \end{aligned}$$

Deux matrices sont alors semblables si, et seulement si, elles ont les mêmes invariants de similitude. Comme les invariants de similitude d'une matrice carrée ne dépendent que de ses coefficients, on obtient en particulier le résultat suivant.

4.9. Corollaire. *Soit K un corps, et soient M, $M' \in \mathrm{M}_n(K)$. Soit L/K une extension de corps. Alors, M et M' sont semblables sur L si, et seulement si, elles sont semblables sur K.*

Autrement dit, s'il existe $P \in \mathrm{GL}_n(L)$ telle que $M' = PMP^{-1}$, alors il existe $P' \in \mathrm{GL}_n(K)$ telle que $M' = P'MP'^{-1}$, ce qui n'est pas a priori immédiat.

Pour achever cette section, nous aimerions donner une démonstration de ce fait lorsque le corps K est infini, dans l'esprit de la démonstration du lemme XXX-5.10. En fait, nous allons même démontrer un résultat un tout petit peu plus général.

4.10. Proposition. *Soient $M_1, M_1', \ldots, M_r, M_r' \in \mathrm{M}_n(K)$, où K est un corps infini. On suppose qu'il existe une extension L/K et $Q \in \mathrm{GL}_n(L)$ telles que*

$$QM_iQ^{-1} = M_i', \quad \textit{pour tout } i \in [\![1, r]\!].$$

Alors, il existe $P \in \mathrm{GL}_n(K)$ telles que

$$PM_iP^{-1} = M_i', \quad \textit{pour tout } i \in [\![1, r]\!].$$

Démonstration. On note E le L-espace vectoriel des matrices $Q \in \mathrm{M}_n(L)$ vérifiant

$$QM_i = M_i'Q, \quad \text{pour tout } i \in [\![1, r]\!].$$

Posons $Q = (x_{ij})_{i,j}$. Alors,

$$Q \in E \iff C \begin{pmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{nn} \end{pmatrix} = 0$$

pour une certaine matrice $C \in \mathrm{M}_{n^2r \times n^2}(K)$.

Notons $V = \{X \in K^{n^2} \mid CX = 0\}$ et $V_L = \{X \in L^{n^2} \mid CX = 0\}$.

Une K-base de V est aussi une L-base de V_L. En effet, si $(X_1, \ldots, X_d)$ est une K-base de V, la matrice dont les colonnes sont $X_1, \ldots, X_d$ est une matrice de $\mathrm{M}_{n^2 \times d}(K)$ de rang d. Mais cette matrice est aussi de rang d, vue comme matrice à coefficients dans L. Ainsi, la famille $(X_1, \ldots, X_d)$, vue comme une famille de V_L, est L-libre. Le rang de C étant le même qu'elle soit considérée comme matrices à coefficients dans K ou coefficients dans L, on en déduit que $\dim_L(V_L) = \dim_K(V)$. Ainsi, $(X_1, \ldots, X_d)$ est une L-base de V_L.

Tout cela montre en particulier que E possède une L-base $(Q_1, \ldots, Q_d)$, avec $Q_i \in \mathrm{M}_n(K)$. Soit $f = \det(X_1Q_1 + \cdots + X_dQ_d) \in K(X_1, \ldots, X_d)$. Alors, f est en fait un élément de $K[X_1, \ldots, X_d]$.

De plus, puisque par hypothèse, il existe au moins un élément de E qui soit inversible, il existe des scalaires $\lambda_1, \ldots, \lambda_d \in L$ tel que $f(\lambda_1, \ldots, \lambda_d) \neq 0$. Ainsi, $f \neq 0$. Comme K est infini, le polynôme f ne peut s'annuler[6] sur tout K^d. Il existe donc $\lambda_1', \ldots, \lambda_d' \in K$ tel que $f(\lambda_1', \ldots, \lambda_d') \neq 0$. Mais alors, la matrice $P = \lambda_1'Q_1 + \cdots + \lambda_d'Q_d$ est inversible à coefficients dans K. Comme $P \in E$, cette matrice convient. Cela achève la démonstration. □

6. Cela se démontre par récurrence sur d.

5. Commutant et bicommutant d'un endomorphisme

Dans tout ce qui suit, E sera un K-espace vectoriel de dimension $n \geqslant 1$ et u sera un endomorphisme de E. Dans cette section finale, nous allons appliquer la décomposition de Frobenius à l'étude du commutant et du bicommutant d'un endomorphisme $u \in \mathscr{L}(E)$.

On commence par définir les objets que nous allons étudier.

5.1. Définition. Soit $u \in \mathscr{L}(E)$. Le *commutant* de u est l'ensemble $C(u)$ des endomorphismes de E qui commutent avec u, c'est-à-dire

$$C(u) = \{u' \in \mathscr{L}(E) \mid u \circ u' = u' \circ u\}.$$

Le *bicommutant* de u est l'ensemble $C_2(u)$ des endomorphismes de E qui commutent avec les éléments de $C(u)$. Autrement dit, $\varphi \in C_2(u)$ si, et seulement si, pour tout $u' \in \mathscr{L}(E)$, on a

$$u \circ u' = u' \circ u \quad \Longrightarrow \quad u' \circ \varphi = \varphi \circ u'.$$

Il est aisé de vérifier que les ensembles $C(u)$ et $C_2(u)$ sont des sous-algèbres de $\mathscr{L}(E)$, qui contiennent $K[u]$ comme sous-algèbre.

L'égalité $C(u) = K[u]$ n'a en revanche pas nécessairement lieu, comme le montre l'exemple de l'endomorphisme identité dès que $n \geqslant 2$. En fait, on sait caractériser les endomorphismes pour lesquels cette égalité est vraie.

5.2. Proposition. *On a $C(u) = K[u]$ si, et seulement si, u est cyclique.*

Démonstration. Supposons u cyclique, et soit $v \in E$ un vecteur u-cyclique. Ainsi, $\big(v,\ u(v),\ \ldots,\ u^{n-1}(v)\big)$ est une base de E.

Soit $u' \in C(u)$, et écrivons

$$u'(v) = a_0 v + a_1 u(v) + \cdots + a_{n-1} u^{n-1}(v), \quad a_i \in K.$$

Posons $P = a_{n-1}X^{n-1} + \cdots + a_1 X + a_0 \in K[X]$, et montrons que $u' = P(u)$.

Par définition, on a $u'(v) = P(u)(v)$. Comme u' commute avec u, il commute avec tout polynôme en u, si bien que pour tout $k \in [\![0, n-1]\!]$, on a

$$u'\big(u^k(v)\big) = u^k\big(u'(v)\big) = u^k\big(P(u)(v)\big) = P(u)\big(u^k(v)\big).$$

Ainsi, u' et $P(u)$ coïncident sur une base, donc sont égaux. On a donc montré que $C(u) \subset K[u]$. L'autre inclusion étant claire, on a $C(u) = K[u]$.

Supposons maintenant que u ne soit pas cyclique, et soit

$$E = F_1 \oplus \cdots \oplus F_r$$

une décomposition de Frobenius de u. Puisque l'endomorphisme u n'est pas cyclique, on a $r \geqslant 2$. Soit $u' \in \mathscr{L}(E)$ la projection sur F_1 parallèlement au sous-espace $F_2 \oplus \cdots \oplus F_r$. Alors, on vérifie facilement que $u' \in C(u)$. En revanche, $u' \notin K[u]$. Supposons le contraire, et écrivons

$$u' = P(u), \quad P \in K[X].$$

Comme u' est nul sur F_2, on a

$$P(u)_{F_2} = P(u_{F_2}) = 0,$$

et donc $\chi_{u,2} \mid P$, puisque $\chi_{u,2}$ est le polynôme minimal de u_{F_2}. Puisque u' est l'identité sur F_1, le même genre de raisonnement montre que $\chi_{u,1} \mid P-1$. Mais, $\chi_{u,1} \mid \chi_{u,2} \mid P$, d'où $\chi_{u,1} \mid \big(P-(P-1)\big)$. En particulier, $\chi_{u,1}$ est constant, d'où une contradiction. Cela achève la démonstration. □

Le lemme suivant sera utile pour l'étude du commutant et du bicommutant.

5.3. Lemme. *Soit $u \in \mathscr{L}(E)$, soient $\chi_{u,1}, \ldots, \chi_{u,r} \in K[X]$ les invariants de similitude de u, et soit $E = F_1 \oplus \cdots \oplus F_r$ une décomposition de Frobenius de u. Enfin, pour tout $i \in [\![1, r]\!]$, soit $v_i \in F_i$ un vecteur u_{F_i}-cyclique.*

Alors, l'application

$$\begin{aligned} \theta : C(u) &\longrightarrow \operatorname{Ker}\big(\chi_{u,1}(u)\big) \times \cdots \times \operatorname{Ker}\big(\chi_{u,r}(u)\big) \\ u' &\longmapsto \big(u'(v_1), \ldots, u'(v_r)\big) \end{aligned}$$

est un isomorphisme de K-espaces vectoriels.

Démonstration. Soit $u' \in C(u)$. Alors, pour tout $i \in [\![1, r]\!]$, et en tenant compte du fait que $v_i \in F_i$, on a

$$\chi_{u,i}(u)\big(u'(v_i)\big) = u'\big(\chi_{u,i}(u)(v_i)\big) = u'\big(\chi_{u,i}(u_{F_i})(v_i)\big) = 0,$$

car $\chi_{u,i}$ le polynôme minimal de u_{F_i}.

Ainsi, $u'(v_i) \in \operatorname{Ker}\big(\chi_{u,1}(u)\big)$ pour tout $i \in [\![1, r]\!]$.

Il est clair que θ est K-linéaire.

Montrons qu'elle est bijective. Soit $u' \in C(u)$ tel que

$$u'(v_1) = \cdots = u'(v_r) = 0.$$

Comme u' commute avec u, on en déduit que aisément que u' s'annule sur la base de Frobenius

$$\big(v_1,\ u(v_1),\ \ldots,\ u^{d_1-1}(v_1),\ \ldots,\quad v_r,\ u(v_r),\ \ldots,\ u^{d_r-1}(v_r)\big),$$

où on a noté $d_i = \deg(\chi_{u,i})$. On en déduit que $u' = 0$, ce qui montre que θ est injective.

Montrons la surjectivité de θ. Soit $y_i \in \operatorname{Ker}\big(\chi_{u,i}(u)\big)$, et soit $u' \in \mathscr{L}(E)$ l'unique endomorphisme de E tel que

$$u'\big(u^k(v_i)\big) = u^k(y_i), \qquad \text{pour tout } i \in [\![1,r]\!] \text{ et tout } k \in [\![0, d_i - 1]\!].$$

Par définition, on a $u'(v_i) = y_i$ pour tout $i \in [\![1,r]\!]$. Vérifions que u' commute avec u. On aura alors $\theta(u') = (y_1, \ldots, y_r)$.

Pour démontrer que $u' \circ u = u \circ u'$, il suffit de vérifier que ces deux endomorphismes coïncident sur la base de Frobenius précédente.

Remarquons que pour tout $P \in K[X]$, et tout $i \in [\![1,r]\!]$, on a

$$u'\big(P(u)(v_i)\big) = P(u)(y_i).$$

En effet, écrivons $P = \chi_{u,i}Q + R$, $Q, R \in K[X]$, avec $\deg(R) \leqslant d_i - 1$. Puisque $\chi_{u,i}$ est le polynôme minimal de u_{F_i}, on a $\chi_{u,i}(u)(v_i) = 0$, et par conséquent, $P(u)(v_i) = R(u)(v_i)$. Mais, par définition de u' et par linéarité, on a

$$u'\big(R(u)(v_i)\big) = R(u)(y_i),$$

puisque R est de degré $\leqslant d_i - 1$.

Par ailleurs, l'hypothèse sur y_i montre que l'on a aussi $P(u)(y_i) = R(u)(y_i)$, d'où les égalités

$$u'\big(P(u)(v_i)\big) = u'\big(R(u)(v_i)\big) = R(u)(y_i) = P(u)(y_i).$$

Mais alors, pour tout $i \in [\![1,r]\!]$ et tout $k \in [\![0, d_i - 1]\!]$, on a

$$(u' \circ u)\big(u^k(v_i)\big) = u'\big(u^{k+1}(v_i)\big) = u^{k+1}(y_i),$$

d'après le point précédent. par ailleurs, on a aussi

$$(u \circ u')\big(u^k(v_i)\big) = u\Big(u'\big(u^k(v_i)\big)\Big) = u\big(u^k(y_i)\big) = u^{k+1}(y_i).$$

Ainsi, les endomorphismes $u' \circ u$ et $u \circ u'$ coïncident sur une base, donc sont égaux. Cela achève la démonstration. □

Nous allons commencer par étudier le bicommutant de u.

5.4. Théorème. *Pour tout endomorphisme $u \in \mathscr{L}(E)$, le bicommutant de u est égal à $K[u]$.*

Démonstration. Soit $E = F_1 \oplus \cdots \oplus F_r$ une décomposition de Frobenius. Pour tout $i \in [\![1,r]\!]$, soit π_i le projecteur sur F_i parallèlement à $\bigoplus_{j \neq i} F_j$.

Soit maintenant $\varphi \in C_2(u)$. Comme F_i et $\bigoplus_{j \neq i} F_j$ sont stables par u, π_i commute avec u, et donc φ commute avec chaque π_i. En particulier, $F_1, \ldots, F_r$ sont stables par φ, et φ_{F_i} commute avec u_{F_i}. Puisque u_{F_i} est cyclique, il

existe un polynôme $Q_i \in K[X]$ tel que

$$\varphi_{F_i} = Q_i(u_{F_i}) = Q_i(u)_{F_i},$$

et cela d'après la proposition 5.2.

Pour tout entier $j \in [\![1, r]\!]$, soit $v_j \in F_j$ un vecteur u_{F_j}-cyclique. Puisque l'on a $\chi_{u,r} = \mu_u$, $\mathrm{Ker}\big(\chi_{u,r}(u)\big) = E$. En particulier, le lemme 5.3 montre qu'il existe $u'_j \in C(u)$ tel que

$$u'_j(v_1) = 0, \ \ldots, \ u'_j(v_{r-1}) = 0, \ u'_j(v_r) = v_j.$$

On a alors d'une part

$$(\varphi \circ u'_j)(v_r) = \varphi\big(u'_j(v_r)\big) = \varphi(v_j),$$

et d'autre part

$$\begin{aligned}(u'_j \circ \varphi)(v_r) &= u'_j\big(\varphi(v_r)\big)\\ &= u'_j\big(\varphi_{F_r}(v_r)\big)\\ &= u'_j\big(Q_r(u)(v_r)\big)\\ &= Q_r(u)\big(u'_j(v_r)\big)\\ &= Q_r(u)(v_j).\end{aligned}$$

Ainsi, on a $\varphi(v_j) = Q_r(u)(v_j)$, pour tout $j \in [\![1, r]\!]$. L'argument usuel nous donne alors l'égalité $\varphi = Q_r(u)$. Ainsi, on a montré que $C_2(u) \subset K[u]$. L'autre inclusion étant déjà connue, cela achève la démonstration. □

On finit ce chapitre en déterminant la dimension de $C(u)$.

5.5. Théorème. *Soient $d_1, \ldots, d_r$ les degrés respectifs des invariants de similitude de u. Alors, on a*

$$\dim_K\big(C(u)\big) = \sum_{k=1}^{r}(2r - 2k + 1)d_k.$$

En particulier, $\dim_K\big(C(u)\big)$ *et* $\dim_K\big(\mathscr{L}(E)\big)$ *ont la même parité.*

Démonstration. D'après le lemme 5.3, dont on garde les notations, on a l'égalité

$$\dim_K\big(C(u)\big) = \sum_{i=1}^{r} \dim_K\Big(\mathrm{Ker}\big(\chi_{u,i}(u)\big)\Big).$$

Soit $x \in E$. On peut donc écrire

$$x = Q_1(u)(v_1) + \cdots + Q_r(u)(v_r),$$

où $\deg(Q_i) < d_i$, et cela de manière unique (il suffit d'écrire x dans la base de Frobenius pour le voir). Alors, $x \in \mathrm{Ker}\big(\chi_{u,i}(u)\big)$ si, et seulement si,

$$(\chi_{u,i}Q_1)(u)(v_1) + \cdots + (\chi_{u,i}Q_r)(u)(v_r) = 0.$$

D'après le corollaire 4.3, cela équivaut à

$$\chi_{u,j} \mid \chi_{u,i} Q_j, \quad \text{pour tout } j \in [\![1, r]\!].$$

Si $j > i$, l'ensemble des polynômes Q_j de degré $< d_j$ vérifiant $\chi_{u,j} \mid \chi_{u,i} Q_j$ sont exactement les polynômes de la forme

$$\frac{\chi_{u,j}}{\chi_{u,i}} R_j, \quad \deg(R_j) < d_i.$$

Si $j \leqslant i$, on a $\chi_{u,j} \mid \chi_{u,i}$, et donc la condition $\chi_{u,j} \mid \chi_{u,i} Q_j$ est automatiquement vérifiée.

On en déduit l'égalité

$$\dim_K \big(\mathrm{Ker}(\chi_{u,i}(u))\big) = \sum_{j=i+1}^{r} d_i + \sum_{j=1}^{i} d_j = (r-i)d_i + \sum_{j=1}^{i} d_j.$$

On a ainsi

$$\begin{aligned}
\dim_K \big(C(u)\big) &= \sum_{i=1}^{r} (r-i)d_i + \sum_{i=1}^{r} \sum_{j=1}^{i} d_j \\
&= \sum_{i=1}^{r} (r-i)d_i + \sum_{j=1}^{r} \sum_{i=j}^{r} d_j \\
&= \sum_{i=1}^{r} (r-i)d_i + \sum_{j=1}^{r} (r-j+1)d_j \\
&= \sum_{k=1}^{r} (2r-2k+1)d_k.
\end{aligned}$$

En particulier, on a

$$\begin{aligned}
\dim_K \big(\mathscr{L}(E)\big) - \dim_K \big(C(u)\big) &= n^2 - \sum_{k=1}^{r} (2r-2k+1)d_k \\
&= n^2 - \sum_{k=1}^{r} (2r-2k)d_k - \sum_{k=1}^{r} d_k \\
&= 2\Big(\frac{n(n-1)}{2} - \sum_{k=1}^{r} (r-k)d_k\Big).
\end{aligned}$$

Cela achève la démonstration. □

5.6. Remarque. Achevons ce chapitre en comparant les avantages et inconvénients des décompositions de Jordan et de Frobenius.

Tout d'abord, elles répondent toutes les deux au problème de classification des endomorphismes/matrices à similitude près. En effet, même si l'existence d'une décomposition de Jordan n'est assurée que si le polynôme caractéristique de l'endomorphisme est scindé, on peut toujours faire les calculs dans une extension L/K convenable (cf. remarque XXX-4.9 (1)), en vertu du corollaire 4.9. En revanche, son calcul nécessite de savoir factoriser le polynôme caractéristique, au moins dans $L[X]$. De plus, dans le

cas où M, $M' \in \mathrm{M}_n(K)$ sont deux matrices semblables, cette astuce ne permettra pas a priori de trouver une matrice de $\mathrm{GL}_n(K)$ qui conjugue M en M', mais seulement une matrice de $\mathrm{GL}_n(L)$ (à moins que les polynômes caractéristiques χ_M et $\chi_{M'}$ ne soient scindés sur K).

En revanche, la décomposition de Frobenius possède l'avantage majeur d'être valable sur tout corps de base, et d'être totalement explicite, sans avoir besoin de savoir factoriser le polynôme caractéristique, et fournit une matrice de passage à coefficients dans K.

La réduction de Frobenius est en général le bon outil pour tous les problèmes liés aux calculs de polynômes minimaux, caractéristiques, aux problèmes de classification, ou de commutation. Elle a en revanche l'inconvénient d'être peu adaptée à tous les problèmes relatifs aux puissances de matrices (calcul explicite, comportement de la suite $(M^n)_{n\geqslant 0}$, calcul d'exponentielles, etc.) ; l'utilisation d'une forme de Jordan est en revanche particulièrement efficace dans ce genre de questions. Enfin, on peut dire que la réduction optimale d'un endomorphisme diagonalisable est une matrice diagonale : c'est le point de vue de la réduction de Jordan et pas celui de la réduction de Frobenius.

6. Exercices

E1. Soit $u \in \mathscr{L}(E)$, où E est de dimension finie non nulle.

On suppose que u est cyclique. Soit $v \in E$ un vecteur u-cyclique.

a. (1) Soit $P \in K[X]$ un diviseur unitaire de μ_u.

Montrer que $F_P = E_{P(u)(v)}$ est stable par u, que u_{F_P} est cyclique, de polynôme minimal $\dfrac{\mu_u}{P}$.

En déduire que si P_1 et P_2 sont deux diviseurs unitaires de μ_u, on a l'équivalence

$$F_{P_1} = F_{P_2} \iff P_1 = P_2.$$

(2) Soit F un sous-espace de E stable par u.

Justifier l'existence d'un polynôme $P \in K[X]$ unitaire de degré minimal tel que $P(u)(v) \in F$.

(3) Montrer que $P(u)(v)$ est un vecteur u_F-cyclique, que $P \mid \mu_u$ et que $\mu_{u_F} = \dfrac{\mu_u}{P}$.

(4) En déduire que l'ensemble des sous-espaces de E stables par u est en bijection avec l'ensemble des diviseurs unitaires de μ_u.

En particulier, E n'a qu'un nombre fini de sous-espaces stables par u.

b. On suppose que K est infini, et que E ne possède qu'un nombre fini de sous-espaces stables par u.

(1) Montrer qu'il existe un vecteur $v \in E$ qui n'appartient à aucun sous-espace stable par u autre que E.

(2) Montrer que v est u-cyclique, et donc que u est un endomorphisme cyclique.

c. On suppose que K est un corps fini. Donner un exemple d'endomorphisme qui admet un nombre fini de sous-espaces stables et qui n'est pas cyclique.

E2. Soit $u \in \mathscr{L}(E)$ un endomorphisme cyclique, et soit $v \in E$ un vecteur u-cyclique. Soit $P \in K[X]$. Donner une condition nécessaire et suffisante sur P pour que le vecteur $P(u)(v)$ soit u-cyclique.

E3. Soit $\pi \in K[X]$ un polynôme irréductible unitaire, et soit $u \in \mathscr{L}(E)$ tel que $\mu_u = \pi$. Quelle est la forme de Frobenius de u ?

E4. Soit

$$M = \begin{pmatrix} 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \in \mathrm{M}_4(\mathbb{Q}),$$

et soit $f : \mathbb{Q}^4 \longrightarrow \mathbb{Q}^4$ l'endomorphisme de $\mathbb{Q}^4$ associé.

Calculer une décomposition de Frobenius de f, ainsi qu'une base de Frobenius.

E5. Montrer que deux matrices de $\mathrm{M}_3(K)$ sont semblables si, et seulement si, elles ont même polynôme minimal et même polynôme caractéristique.

Est-ce vrai en dimension 2 ? en dimension $n \geqslant 4$?

E6. Soit K un corps, et soit $M \in \mathrm{M}_n(K)$.

a. Montrer sans calculs que M et M^t sont semblables.

b. En considérant la matrice

$$M = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \in \mathrm{M}_3(\mathbb{Z}),$$

montrer que le résultat de **a** n'est plus vrai si l'on remplace K par un anneau.

c. Soit $M \in \mathrm{M}_n(K)$ une *matrice compagnon*. En traduisant en termes de coefficients, montrer que toute solution $S \in \mathrm{M}_n(K)$ de l'équation

$$M^t S = SM,\ S \in \mathrm{M}_n(K)$$

est symétrique. En utilisant la question **a**, montrer l'existence d'une solution S inversible (nécessairement symétrique).

d. Soit $M \in \mathrm{M}_n(K)$ une matrice quelconque. Montrer qu'il existe une matrice inversible *symétrique* $S \in \mathrm{GL}_n(K)$ telle que $SMS^{-1} = M^t$.

e. En déduire que toute matrice de $\mathrm{M}_n(K)$ est produit de deux matrices symétriques.

f. On se propose de retrouver le résultat de la question **c** de manière plus conceptuelle.

Soit $P = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0 \in K[X]$ un polynôme unitaire de degré $n \geqslant 1$, et soit $E = K[X]/(P)$.

(i) Montrer que la matrice représentative de multiplication par $\overline{X}$ dans la base $(\overline{1}, \overline{X}, \ldots, \overline{X^{n-1}})$ est C_P.

(ii) On définit par récurrence des polynômes $H_0, \ldots, H_{n-1}$ par

$$H_0 = 1 \quad \text{et} \quad H_k = XH_{k-1} + a_{n-k} \quad \text{pour tout } k \in [\![0, n-1]\!].$$

Vérifier que $(\overline{H}_{n-1}, \ldots, \overline{H}_0)$ est une base de E, que la matrice représentative de la multiplication par $\overline{X}$ dans E dans cette nouvelle base est C_P^t, et que la matrice de passage est symétrique.

(iii) Retrouver le résultat de la question **c**.

E7. Soit $u \in \mathscr{L}(E)$, où E est de dimension $n \geqslant 1$.

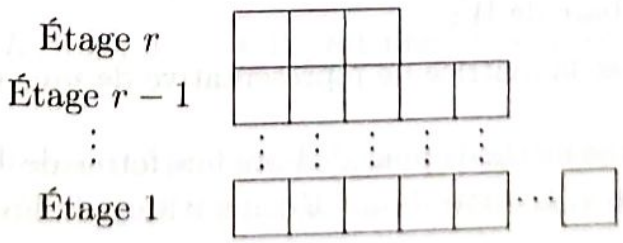
Montrer que $\dim_K \big(C(u)\big) \geqslant n$, avec égalité si, et seulement si, u est cyclique.

E8. Soit $n \geqslant 1$ un entier.

a. Une *partition* de n est une suite d'entiers $(p_k)_{k\geqslant 1}$ décroissante, nulle à partir d'un certain rang, telle que $\sum_{k\geqslant 1} p_k = n$.

Si $(p_k)_{k\geqslant 1}$ est une partition de n, on lui associe un tableau, dit tableau de Young.

On procède ligne par ligne :



où la ligne k possède p_k cases.

Inversement, un tel tableau de Young (où le nombre de cases sur chaque ligne décroît et stationne à 0) à n cases correspond à une unique partition de n. Pour $j \geqslant 1$, on note p_j^* le nombre de cases sur la colonne j.

(1) Montrer que $(p_j^*)_{j\geqslant 1}$ est une partition de n.

Indication. On pourra montrer que $p_j^* = |\{k \geqslant 1 \mid p_k \geqslant j\}|$.

(2) Montrer l'égalité

$$\sum_{k\geqslant 1}(2k-1)p_k = \sum_{j\geqslant 1} p_j^{*2}.$$

Indication. On suppose l'égalité voulue démontrée pour une partition $(p_k)_{\geqslant 1}$. Montrer que si l'on rajoute une case sur le tableau de Young correspondant, de façon à obtenir encore un tableau de Young, la nouvelle partition obtenue vérifie encore l'égalité.

b. Soit K un corps, et soit E un K-espace vectoriel de dimension $n \geqslant 1$. Soit $m \geqslant 1$ un entier. Montrer qu'il existe $u \in \mathscr{L}(E)$ tel que

$$\dim_K\big(C(u)\big) = m$$

si, et seulement si, il existe des entiers $q_1, \ldots, q_s \geqslant 1$ tels que

(i) $q_1 + \cdots + q_s = n$

(ii) $q_1^2 + \cdots + q_s^2 = m$.

E9. Soit E un K-espace vectoriel de dimension $n \geqslant 1$, et soit $u \in \mathscr{L}(E)$ de polynôme caractéristique scindé sur K.

a. On suppose que u est cyclique. Pour tout $\lambda \in \mathrm{Sp}_K(u)$, on pose

$$Q_\lambda = \frac{\chi_u}{(X-\lambda)^{m_\lambda}},$$

où m_λ est la multiplicité de λ dans χ_u.

Si $v \in E$ est un vecteur u-cyclique, on pose enfin $w_\lambda = Q_\lambda(u)(v)$.

(1) Montrer que le sous-espace cyclique W_λ engendré par w_λ est stable par u, de dimension m_λ, que le polynôme minimal de u_{W_λ} est égal à $(X-\lambda)^{m_\lambda}$, et que la famille

$$\big((u-\lambda\mathrm{Id}_{W_\lambda})^{m_\lambda-1}(w_\lambda),\ \ldots,\ (u-\lambda\mathrm{Id}_{W_\lambda})(w_\lambda),\ w_\lambda\big)$$

est une base de W_λ.

(2) Quelle est la matrice de représentative de u_{W_λ} dans cette base ?

b. En déduire une méthode pour déduire une forme de Jordan et une base de Jordan de u, à partir d'une décomposition de Frobenius.

c. Calculer une forme de Jordan, ainsi qu'une base de Jordan, de l'endomorphisme $u \in \mathscr{L}(\mathbb{Q}^5)$ dont la matrice dans la base canonique est

$$M = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

E10. Soit E un K-espace vectoriel de dimension $n \geqslant 1$, et soit $u \in \mathscr{L}(E)$ de polynôme caractéristique scindé sur K.

a. On suppose que $E = F_1 \oplus \cdots \oplus F_s$, où chaque F_i est stable par u, et possède une base dans laquelle la matrice représentative de u_{W_i} est une cellule de Jordan J_{m_i,λ_i}, $\lambda_1, \ldots, \lambda_s \in K$ sont deux à deux distincts.

On note $w_i \in W_i$ le dernier vecteur de la base correspondante. On pose

$$\chi = (X - \lambda_1)^{m_1} \cdots (X - \lambda_s)^{m_s},$$

ainsi que

$$Q_i = \frac{\chi}{(X - \lambda_i)^{m_i}}, \quad \text{pour tout } i \in [\![1, s]\!].$$

Si $U_1, \ldots, U_s \in K[X]$ vérifient $Q_1U_1 + \cdots + Q_sU_s = 1$, montrer que l'on a $E = E_w$, avec $w = U_1(u)(w_1) + \cdots U_s(u)(w_s)$.

b. On revient au cas général. Pour $\lambda \in \mathrm{Sp}_K(\lambda)$, on décompose le sous-espace caractéristique

$$E_\lambda^{m_\lambda} = \bigoplus_{j=1}^{t_\lambda} W_{j,\lambda},$$

où $W_{j,\lambda}$ est stable par u, et possède une base dans laquelle la matrice représentative de u_{W_i} est une cellule de Jordan. On suppose aussi que l'on a

$$\dim_K(W_{j,\lambda}) \geqslant \dim_K(W_{j+1,\lambda}), \quad \text{pour tout } j \in [\![1, t_\lambda - 1]\!].$$

On pose $t = \max_\lambda t_\lambda$, ainsi que $W_{j,\lambda} = \{0\}$ pour tout $j > t_\lambda$. Enfin, on pose

$$W_j = \bigoplus_{\lambda \in \mathrm{Sp}_K(u)} W_{j,\lambda}, \quad \text{pour tout } j \in [\![1, t]\!].$$

Montrer que $E = W_1 \oplus \cdots \oplus W_t$ est une décomposition de Frobenius de u.

c. Calculer une forme de Frobenius, ainsi qu'une base de Frobenius, de l'endomorphisme $u \in \mathscr{L}(\mathbb{Q}^4)$ dont la matrice dans la base canonique

est

$$M = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & -1 \end{pmatrix}.$$

E11. Donner les invariants de similitude d'une matrice diagonale.

E12. Soit L/K une extension galoisienne de degré $n \geqslant 1$. On suppose que $\mathrm{Gal}(L/K)$ est cyclique, engendré par un élément σ (c'est donc en particulier le cas si K est fini).

a. Utiliser le lemme d'indépendance de Dedekind pour démontrer que le polynôme minimal de σ, vu comme endomorphisme du K-espace vectoriel L, est égal à $X^n - 1$.

b. Montrer que σ est un endomorphisme cyclique.

c. En déduire qu'il existe $x \in L$ tel que $\big(x,\ \sigma(x),\ \ldots,\ \sigma^{n-1}(x)\big)$ soit une K-base de L.

Chapitre XXXII

Dualité

1. Introduction

Soit $A \in \mathrm{M}_{m\times n}(K)$ une matrice de rang r, et soit $AX = 0$, avec $X \in K^n$, le système linéaire correspondant.

Il est bien connu que l'ensemble des solutions d'un tel système linéaire est un sous-espace de K^n de dimension $n-r$. Réciproquement, tout sous-espace de K^n de dimension $n-r$, avec $r \in [\![0, n]\!]$, est l'ensemble des solutions d'un système linéaire associé à une matrice de rang r.

Il y a deux démonstrations de ce résultat un peu différentes, selon que l'on se place dans le cas d'un corps de base quelconque ou sur $\mathbb{R}$.

Supposons tout d'abord que $K = \mathbb{R}$. Posons $B = A^t$. Si W est le sous-espace de $\mathbb{R}^n$ engendré par les colonnes de B (i.e. les lignes de A), alors l'ensemble des solutions du système $AX = 0$ n'est rien d'autre que l'orthogonal $W^\perp$ du sous-espace W pour le produit scalaire usuel sur $\mathbb{R}^n$. Or, la matrice A étant de rang r, B aussi, et donc $\dim_{\mathbb{R}}(W) = r$. Mais alors, $\dim_{\mathbb{R}}(W^\perp) = n - r$.

Inversement, si V est un sous-espace de $\mathbb{R}^n$ de dimension $n - r$, alors $V^\perp$ est de dimension r. Or, on a $V = (V^\perp)^\perp$. Si maintenant $(v'_1, \ldots, v'_r)$ est une base de $V^\perp$, et si B est la matrice dont les colonnes sont $v'_1, \ldots, v'_r$, V est donc l'ensemble des solutions du système linéaire $B^tX = 0$.

Tout cela utilise fortement la structure euclidienne sur $\mathbb{R}^n$. D'ailleurs, les résultats dont dépendent cette démonstration peuvent se résumer ainsi : pour tout $r \in [\![0, n]\!]$, les applications

$$\begin{array}{ccc} V & \longmapsto & V^\perp \\ W^\perp & \longleftarrow & W \end{array}$$

établissent une correspondance bijective entre les sous-espaces V de dimension $n - r$ et les sous-espaces W de dimension r de $\mathbb{R}^n$.

Notons que cela s'interprète bien géométriquement si $n = 3$ et $r = 1$: l'ensemble des vecteurs orthogonaux à une droite vectorielle de $\mathbb{R}^3$ est un plan vectoriel, et toute droite vectorielle de $\mathbb{R}^3$ est l'intersection de deux plans vectoriels.

Supposons maintenant que K soit un corps quelconque. Alors, l'algorithme de Gauss (ou alors le théorème du rang appliqué à l'application linéaire représentée par A) permet de montrer que l'ensemble des solutions du système linéaire

$$AX = 0, \quad X \in K^n$$

est un sous-espace de K^n de dimension $n - r$.

Inversement, soit $r \in [\![0, n]\!]$. Pour démontrer que tout sous-espace de K^n de dimension $n-r$ est l'ensemble des solutions d'un système linéaire associé à une matrice de rang r, on peut procéder ainsi. Soit W un sous-espace de K^n de dimension $n - r$, et soit $(w_1, \ldots, w_{n-r})$ une base de W. Considérons la matrice $C \in \mathrm{M}_{n\times n-r}(K)$ dont les colonnes sont $w_1, \ldots, w_{n-r}$. C'est une matrice de rang $n-r$. Soit $Y \in K^n$. Par définition de C, le vecteur $Y \in W$ si, et seulement si, le système $CX = Y$ a au moins une solution. En appliquant l'algorithme de Gauss, on constate que le système $CX = Y$ a une solution si, et seulement si, Y vérifie certaines conditions de compatibilité s'écrivant linéairement en fonction des coordonnées de Y, c'est-à-dire si, et seulement si, $AY = 0$ pour une certaine matrice $A \in \mathrm{M}_{p\times n}(K)$. Le fait que W soit de dimension $n - r$ implique que $\mathrm{Ker}(A)$ est de dimension $n - r$, i.e. que A de rang r.

Cette « démonstration » peut paraître un peu miraculeuse. Elle ne l'est en fait pas tant que cela. En réinterprétant convenablement les objets, nous allons nous apercevoir qu'il y a beaucoup de points communs aux deux approches.

L'ensemble des solutions d'une équation de la forme

$$a_1x_1 + \cdots + a_nx_n = 0, \quad x_i \in K,\ i \in [\![1, n]\!]$$

n'est rien d'autre que le noyau de la forme linéaire

$$\begin{array}{ccc} K^n & \longrightarrow & K \\ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} & \longmapsto & a_1x_1 + \cdots + a_nx_n, \end{array}$$

et l'ensemble des solutions d'un système linéaire de la forme $AX = 0$ peut se voir comme l'intersection des noyaux de n formes linéaires.

Le résultat sur les systèmes linéaires se retraduit alors comme suit : si E est un K-espace vectoriel de dimension n, et $\varphi_1, \ldots, \varphi_n : K^n \longrightarrow K$ sont

des formes linéaires engendrant un sous-espace de $(K^n)^*$ de dimension r, alors $\bigcap_{i=1}^{n} \mathrm{Ker}(\varphi_i)$ est de dimension [1] égale à $n-r$. De plus, tout sous-espace de dimension $n-r$ peut s'écrire comme l'intersection des noyaux de r formes linéaires indépendantes.

Pour voir la similitude des deux approches, l'idée est, étant donné un espace vectoriel E de dimension n sur K, d'essayer d'établir une correspondance entre les sous-espaces de dimension r du K-espace vectoriel $E^* = \mathscr{L}(E, K)$ et les sous-espaces de dimension $n-r$ de E. Dans le cas de $\mathbb{R}^n$, cette correspondance entre sous-espaces a été établie en utilisant une forme bilinéaire particulière, à savoir le produit scalaire usuel sur $\mathbb{R}^n$. Elle se généralise d'ailleurs à n'importe quel espace euclidien.

Or, pour tout K-espace vectoriel E (pas nécessairement de dimension finie), il existe une forme bilinéaire naturelle $E \times E^* \longrightarrow K$, à savoir

$$\begin{aligned} \langle\ ,\ \rangle_E : E \times E^* &\longrightarrow K \\ (x, \varphi) &\longmapsto \varphi(x). \end{aligned}$$

1.1. Définition. Soit E un K-espace vectoriel (de dimension quelconque). L'application bilinéaire

$$\begin{aligned} \langle\ ,\ \rangle_E : E \times E^* &\longrightarrow K \\ (x, \varphi) &\longmapsto \varphi(x) \end{aligned}$$

s'appelle le *crochet de dualité* sur E.

On peut maintenant comme dans le cas euclidien définir la notion d'orthogonalité [2], en disant que $x \in E$ et $\varphi \in E^*$ sont orthogonaux si $\langle x, \varphi \rangle_E = 0$. Notons au passage que, contrairement au cas euclidien, les deux éléments considérés ne vivent pas dans le même espace vectoriel. Néanmoins, on peut malgré tout définir l'orthogonal d'un sous-espace. En fait ici, on a deux choix possibles, car deux espaces vectoriels différents interviennent.

Si W est un sous-espace de E^*, on définit son orthogonal par

$${}^\circ W = \{x \in E \mid \langle x, \varphi \rangle_E = 0, \quad \text{pour tout } \varphi \in W\}.$$

Autrement dit, on a

$${}^\circ W = \bigcap_{\varphi \in W} \mathrm{Ker}(\varphi).$$

Mais on peut aussi définir l'orthogonal d'un sous-espace V de E par

$$V^\circ = \{\varphi \in E^* \mid \langle x, \varphi \rangle_E = 0, \quad \text{pour tout } x \in V\},$$

c'est-à-dire

$$V^\circ = \{\varphi \in E^* \mid V \subset \mathrm{Ker}(\varphi)\}.$$

Attention ! Cette notation n'est pas standard, et la plupart des ouvrages notent plutôt ces orthogonaux $W^\perp$ et V° respectivement.

1. Cf. lemme III-5.10
2. Même si cela n'a plus rien à voir avec la géométrie...

On s'attend alors à ce que le résultat suivant soit vrai : si E est un K-espace vectoriel de dimension $n \geqslant 1$, alors pour tout $r \in [\![0, n]\!]$, les applications

$$\begin{array}{ccc} V & \longmapsto & V^\circ \\ {}^\circ W & \longmapsfrom & W \end{array}$$

établissent une correspondance bijective entre les sous-espaces de E de dimension $n - r$ et les sous-espaces de E^* dimension r. Autrement dit, on devrait avoir

$$^\circ(V^\circ) = V \quad \text{et} \quad (^\circ W)^\circ = W$$

pour tout sous-espace V de E et tout sous-espace W de E^*.

Remarquons que le cas $r = 1$ est contenu dans l'énoncé du lemme III-5.8 (en notant que $^\circ(K\varphi) = \mathrm{Ker}(\varphi)$).

Ce résultat, s'il est vrai[3], nous permettrait alors de démontrer notre énoncé de manière un peu plus conceptuelle, avec une approche similaire à celle du cas euclidien.

Cela nous encourage donc à essayer d'établir une correspondance entre les sous-espaces de deux K-espaces vectoriels E et F, lorsque qu'il existe une application bilinéaire $b : E \times F \longrightarrow K$. C'est que nous ferons dans la suite.

2. Formes linéaires et bilinéaires, espace dual

Commençons par rappeler la définition d'une application bilinéaire.

2.1. Définition. Soient E, F deux K-espaces vectoriels. Une *forme bilinéaire* est une application

$$\begin{array}{rccc} b : & E \times F & \longrightarrow & K \\ & (x, y) & \longmapsto & b(x, y) \end{array}$$

telle que pour tout $x \in E$, $y \in F$ les applications

$$\begin{array}{rccc} b(x, \_) : & F & \longrightarrow & K \\ & y & \longmapsto & b(x, y) \end{array} \quad \text{et} \quad \begin{array}{rccc} b(\_, y) : & E & \longrightarrow & K \\ & x & \longmapsto & b(x, y) \end{array}$$

soient linéaires, c'est-à-dire une application linéaire en chaque argument.

2.2. Exemples

(1) L'application

$$\begin{array}{rccc} & \mathbb{R}[X] \times \mathbb{R}[X] & \longrightarrow & \mathbb{R} \\ & (P, Q) & \longmapsto & \displaystyle\int_0^1 P(t)Q(t)e^{-t^2}\,\mathrm{d}t \end{array}$$

3. Et il l'est !

est bilinéaire. C'est même un produit scalaire.

(2) Le crochet de dualité

$$\begin{aligned} \langle\ ,\ \rangle_E : E \times E^* &\longrightarrow K \\ (x, \varphi) &\longmapsto \langle x, \varphi\rangle = \varphi(x) \end{aligned}$$

est bilinéaire.

La démonstration du lemme suivant est facile et laissée en exercice au lecteur.

2.3. Lemme. *Soit $b : E \times F \longrightarrow K$ une application bilinéaire. Alors, pour tout $x \in E$, et tout $y \in F$, les applications $b(x,\_)$ et $b(\_, y)$ sont des formes linéaires, et les applications*

$$\begin{aligned} {}_*b : E &\longrightarrow F^* \\ x &\longmapsto b(x,\_) \end{aligned} \qquad \textit{et} \qquad \begin{aligned} b_* : F &\longrightarrow E^* \\ y &\longmapsto b(\_, y) \end{aligned}$$

sont linéaires.

2.4. Exemple. Si $F = E^*$ et $b : E \times E^* \longrightarrow K$ est le crochet de dualité, on a $b_* = \mathrm{Id}_{E^*}$ et ${}_*b$ est l'application $\chi : E \longrightarrow E^{**}$ qui à $x \in E$ associe l'évaluation en x, i.e.

$$\chi(x)(\varphi) = \varphi(x), \qquad \text{pour tout } \varphi \in E^*.$$

Ici, E^{**} désigne l'espace $(E^*)^*$.

De même que l'on a défini une notion d'orthogonalité pour les produits scalaires ou pour le crochet de dualité, on a une notion d'orthogonalité par rapport à une forme bilinéaire $b : E \times F \longrightarrow K$.

2.5. Définition. Soit $b : E \times F \longrightarrow K$ une forme bilinéaire.

Si $x \in E$ et $y \in F$, on dit que x et y sont *orthogonaux* si $b(x, y) = 0$.

On le note parfois $x \perp_b y$ ou $x \perp y$, lorsque il n'y a pas d'ambiguïté sur le choix de b.

Si V est un sous-espace de E, on définit *l'orthogonal* de V par

$$V^\perp = \{y \in F \mid b(x, y) = 0, \ \text{ pour tout } x \in V\}.$$

C'est un sous-espace vectoriel de F.

Si W est un sous-espace de F, *l'orthogonal* de W est

$${}^\perp W = \{x \in E \mid b(x, y) = 0, \ \text{ pour tout } y \in W\}.$$

C'est un sous-espace vectoriel de E.

Attention ! La formulation « x et y orthogonaux » semble supposer que la relation d'orthogonalité est symétrique. Il n'en est rien, puisqu'a priori les vecteurs x et y ne vivent pas dans le même espace vectoriel. Cela peut être très piégeux lorsque $E = F$. En effet, si $b : E \times F \longrightarrow K$ est une application bilinéaire quelconque, en général, $x \perp_b y$ n'implique **PAS** que $y \perp_b x$.

On peut d'ailleurs démontrer que si $b : E \times E \longrightarrow K$ est bilinéaire, alors la relation d'orthogonalité (par rapport à b) est symétrique si, et seulement si, b est symétrique ou antisymétrique (cf. exercice 3).

2.6. Remarques

(1) Si $F = E^*$ et $b = \langle\ ,\ \rangle_E$ est le crochet de dualité, ces notions coïncident avec celles définies dans la section précédente, à la notation près.

(2) Lorsque $E = F$ et b est symétrique (par exemple si E est un espace vectoriel réel et b est un produit scalaire), on a $b_* = {}_*b$, et les deux définitions d'orthogonal d'un sous-espace coïncident.

(3) Si $b : E \times F \longrightarrow K$ est une application bilinéaire, les définitions montrent que l'on a

$$V \subset {}^\perp(V^\perp) \quad \text{et} \quad W \subset ({}^\perp W)^\perp,$$

pour tout sous-espace V de E et tout sous-espace W de F.

En effet, soit $x \in V$. Pour tout $y \in V^\perp$, on a $b(x, y) = 0$, et par conséquent $x \in {}^\perp(V^\perp)$. L'autre inclusion se montre de même.

Dans ce qui suit, on considère deux K-espaces vectoriels E et F, et une application K-bilinéaire

$$b : E \times F \longrightarrow K.$$

Sauf mention expresse du contraire, E et F seront non nuls.

Rappelons que notre but est de comprendre à quelles conditions les applications

$$\begin{array}{ccc} V & \longmapsto & V^\perp \\ {}^\perp W & \longmapsfrom & W \end{array}$$

sont inverses l'une de l'autre, établissant ainsi une correspondance bijective entre les sous-espaces de E et les sous-espaces de F.

Lorsque c'est le cas, on a

$${}^\perp(V^\perp) = V \quad \text{et} \quad ({}^\perp W)^\perp = W,$$

pour tout sous-espace V de E et tout sous-espace W de F. En particulier, on doit avoir

$${}^\perp(\{0_E\}^\perp) = \{0_E\} \quad \text{et} \quad ({}^\perp\{0_F\})^\perp = \{0_F\}.$$

Puisque $b(0_E, y) = 0$ pour tout $y \in F$, on a $\{0_E\}^\perp = F$, et ainsi

$$\begin{aligned} {}^\perp(\{0_E\}^\perp) &= {}^\perp F \\ &= \{x \in E \mid b(x,y) = 0, \text{ pour tout } y \in F\} \\ &= \{x \in E \mid b(x,\_) = 0\} \\ &= \mathrm{Ker}({}_*b). \end{aligned}$$

De même, on a

$$({}^\perp\{0_F\})^\perp = \mathrm{Ker}(b_*).$$

Ainsi, une condition nécessaire pour avoir cette correspondance [4] est l'injectivité des deux applications ${}_*b$ et b_*. Cela conduit à la définition suivante.

2.7. Définition. On dit qu'une forme bilinéaire $b : E \times F \longrightarrow K$ est *non dégénérée à gauche* si l'application ${}_*b : E \longrightarrow F^*$ est injective. Autrement dit, $b : E \times F \longrightarrow K$ est non dégénérée à gauche si l'on a

$$\text{pour tout } x \in E,\ b(x,y) = 0,\ \text{pour tout } y \in F \quad \Longrightarrow \quad x = 0_E.$$

De manière équivalente, $b : E \times F \longrightarrow K$ est non dégénérée à gauche si pour tout $x \in E \setminus \{0_E\}$, il existe $y \in F$ tel que $b(x,y) \neq 0$.

On dit qu'une forme bilinéaire $b : E \times F \longrightarrow K$ est *non dégénérée à droite* si l'application $b_* : F \longrightarrow E^*$ est injective. Autrement dit, b est non dégénérée à droite si l'on a

$$\text{pour tout } y \in F,\ b(x,y) = 0,\ \text{pour tout } x \in E \quad \Longrightarrow \quad y = 0_F.$$

De manière équivalente, $b : E \times F \longrightarrow K$ est non dégénérée à droite si pour tout $y \in F \setminus \{0_F\}$, il existe $x \in F$ tel que $b(x,y) \neq 0$.

On dit qu'une forme bilinéaire $b : E \times F \longrightarrow K$ est *non dégénérée* si elle est non dégénérée à gauche et à droite.

Tout produit scalaire sur un espace vectoriel réel est bien entendu une forme bilinéaire non dégénérée. Voici un autre exemple.

2.8. Exemple. Soit E un K-espace vectoriel. Alors, le crochet de dualité est une forme bilinéaire non dégénérée.

En effet, comme on l'a déjà remarqué plus haut, l'application $(\langle\ ,\ \rangle_E)_*$ est simplement Id_{E^*}, qui est injective. Le crochet de dualité est donc non dégénéré à droite.

Soit maintenant $x \neq 0_E$. Nous allons construire une forme linéaire $\varphi_0 \in E^*$ telle que $\varphi_0(x) = 1$. Pour cela, choisissons un supplémentaire [5] V' de Kx

4. Via les deux applications décrites ci-dessus.
5. L'existence d'un tel supplémentaire est assurée par la proposition III-3.5.

dans E. Posons

$$\varphi_0(\lambda x + v') = \lambda, \quad \text{pour tout } \lambda \in K \text{ et tout } v' \in V'.$$

Alors, $\varphi_0 \in E^*$ et $\varphi_0(x) = 1$.

Cela montre en particulier que le seul vecteur $x \in E$ tel que

$$\varphi(x) = 0, \quad \text{pour tout } \varphi \in E^*$$

est le vecteur nul. Ainsi, le crochet de dualité est aussi non dégénéré à gauche.

2.9. Remarque. Soit $b : E \times F \longrightarrow K$ une forme bilinéaire non dégénérée. Si l'un des deux espaces E ou F est de dimension finie, alors il en est de même de l'autre. Dans ce cas, on a $\dim_K(E) = \dim_K(F)$, et les deux applications linéaires ${}_*b$ et b_* sont des isomorphismes.

En effet, supposons par exemple que E soit de dimension finie sur K. Alors, E^* est aussi de dimension finie. Puisque $b_* : F \longrightarrow E^*$ est injective, F s'identifie à un sous-espace de E^* et est donc de dimension finie. De plus, on a

$$\dim_K(F) \leqslant \dim_K(E^*) = \dim_K(E).$$

Comme ${}_*b : E \longrightarrow F^*$ est injective, on a

$$\dim_K(E) \leqslant \dim_K(F^*) = \dim_K(F).$$

Ainsi, $\dim_K(E) = \dim_K(F)$. Les K-espaces vectoriels E, E^*, F et F^* ont alors même dimension, ce qui implique que ${}_*b$ et b_* sont des isomorphismes (puisqu'elle sont injectives).

Ce résultat ne subsiste plus en dimension infinie, comme le montre l'exemple suivant.

2.10. Exemple. Soit E un K-espace vectoriel, et soit $b = \langle\ ,\ \rangle_E$ le crochet de dualité. L'application ${}_*b$ est donc l'application $\chi : E \longrightarrow E^{**}$ qui à x associe l'évaluation en x.

On a alors les équivalences

$$\chi \text{ surjective} \iff \chi \text{ bijective} \iff \dim_K(E) < +\infty.$$

La première équivalence est claire, puisque χ est injective. Il reste à démontrer la seconde.

Si E est de dimension finie, on a

$$\dim_K(E) = \dim_K(E^*) = \dim_K(E^{**}),$$

et comme χ est injective, elle est aussi surjective.

Avant de voir ce qui se passe en dimension infinie, essayons de comprendre un peu mieux l'image de χ.

Soit $\mathbf{e} = (e_i)_{i\in I}$ une base de E. On note $(e_i^*)_{i\in I}$ la famille des formes coordonnées associées à la base $\mathbf{e}$ (cf. définition III-5.3). Cette famille est libre par la proposition III-5.4.

Pour tout $x \in E$, on écrit $x = \sum_{i\in I} x_i e_i$, où les $x_i \in K$ sont presque tous nuls. Il existe donc un sous-ensemble fini J de I tel que $x_i = 0$ pour tout indice $i \in I \setminus J$. On a alors

$$\chi(x)(e_i^*) = e_i^*(x) = x_i = 0, \quad \text{pour tout } i \in I \setminus J.$$

Si maintenant E est de dimension infinie, complétons la famille libre $(e_i^*)_{i\in I}$ en une base $(\varphi_k)_{k\in I'}$. Soit $\rho : E^* \longrightarrow K$ l'unique forme linéaire vérifiant

$$\rho(\varphi_k) = 1, \quad \text{pour tout } k \in I'.$$

Alors, comme la base $(\varphi_k)_{k\in I'}$ contient la famille $(e_i^*)_{i\in I}$, on a

$$\rho(e_i^*) = 1 \neq 0, \quad \text{pour tout } i \in I.$$

Au vu de ce qui précède, on en déduit que $\rho \notin \mathrm{Im}(\chi)$.

Le lemme suivant montre que la bijectivité des deux applications b_* et ${}_*b$ ne peut être obtenue qu'en dimension finie.

2.11. Lemme. *Soit $b : E \times F \longrightarrow K$ une forme bilinéaire. Alors :*

(1) *si b_* est surjective, ${}_*b$ est injective ;*

(2) *si ${}_*b$ est surjective, b_* est injective ;*

(3) *les propriétés suivantes sont équivalentes :*

(i) *les applications linéaires b_* et ${}_*b$ sont bijectives ;*

(ii) *la forme bilinéaire b est non dégénérée, et E et F sont de dimension finie.*

Dans ce cas, $\dim_K(E) = \dim_K(F)$.

Démonstration. Supposons que b_* soit surjective, et soit $x \in E$ un vecteur de E tel que ${}_*b(x) = 0$. Alors, pour tout $y \in F$, on a

$$b_*(y)(x) = b(x, y) = {}_*b(x)(y) = 0.$$

Comme b_* est surjective, on en déduit que $\varphi(x) = 0$ pour tout $\varphi \in E^*$, ce qui implique que $x = 0_E$ par l'exemple 2.8, d'où (1). Le point (2) se démontre de même.

Si b est non dégénérée, et E et F sont de dimension finie, alors b_* et ${}_*b$ sont bijectives et $\dim_K(E) = \dim_K(F)$, d'après la remarque 2.9.

Inversement, supposons que b_* et ${}_*b$ soient bijectives. Elles sont donc injectives et b est non dégénérée. Nous allons maintenant montrer que E est de dimension finie, ce qui suffira pour conclure par la remarque 2.9.

On rappelle que l'application $\chi : E \longrightarrow E^{**}$ associe à $x \in E$ l'évaluation en x, c'est-à-dire

$$\chi(x)(\varphi) = \varphi(x), \quad \text{pour tout } \varphi \in E^*.$$

Considérons l'application

$$\begin{array}{rcl} r_{b_*} : E^{**} & \longrightarrow & F^* \\ \rho & \longmapsto & \rho \circ b_*. \end{array}$$

Elle est linéaire, et puisque b_* est bijective, elle est aussi bijective, d'inverse

$$\begin{array}{rcl} F^* & \longrightarrow & E^{**} \\ \psi & \longmapsto & \psi \circ b_*^{-1}. \end{array}$$

Remarquons maintenant que pour tout $y \in F$, et tout $x \in E$, on a

$$\begin{array}{rcl} \big(r_{b_*}(\chi(x))\big)(y) & = & \big(\chi(x) \circ b_*\big)(y) \\ & = & \chi(x)\big(b_*(y)\big) \\ & = & b_*(y)(x) \\ & = & b(x,y) \\ & = & {}_*b(x)(y). \end{array}$$

Ainsi, pour tout $x \in E$, on a $r_{b_*}\big(\chi(x)\big) = {}_*b(x)$, d'où l'égalité

$$r_{b_*} \circ \chi = {}_*b.$$

Comme ${}_*b$ et r_{b_*} sont bijectives, $\chi : E \longrightarrow E^{**}$ est bijective, et E est donc de dimension finie d'après l'exemple 2.10. Cela achève la démonstration. □

Nous continuons cette section en remarquant que la correspondance bien connue entre sous-espaces de E dans le cas euclidien, ou entre les sous-espaces de E et ceux de E^* lorsque E est de dimension finie[6] n'ont plus cours en dimension infinie.

2.12. Exemples

(1) Soit E l'espace des suites réelles $\mathbf{u} = (u_n)_{n \geqslant 0}$ telles que la série

$$\sum_{n \geqslant 0} u_n^2$$

converge, muni du produit scalaire

$$\langle \mathbf{u}, \mathbf{v} \rangle = \sum_{n \geqslant 0} u_n v_n,$$

6. Correspondance pas encore tout à fait démontrée, mais fortement suggérée par son interprétation en termes de systèmes linéaires.

et soit V le sous-espace des suites nulles à partir d'un certain rang. Alors $V^\perp = \{0\}$, et donc ${}^\perp(V^\perp) = E \neq V$.
En effet, V contient $\mathbf{e}_j = (\delta_{nj})_{n\geqslant 0}$, pour tout entier $j \geqslant 0$. Or, pour tout $j \geqslant 0$, on a

$$\langle \mathbf{u}, \mathbf{e}_j \rangle = u_j.$$

En particulier, une suite orthogonale à tout élément de V est forcément la suite nulle.

(2) Soit E un K-espace vectoriel, de base $\mathbf{e} = (e_i)_{i\in I}$. Supposons que E soit de dimension infinie et soit $W = \mathrm{Vect}_K(e_i^*,\ i \in I)$.
Par la proposition III-5.4, $W \neq E^*$. En revanche, on a ${}^\circ W = \{0_E\}$, et ainsi $({}^\circ W)^\circ = E^* \neq W$. En effet, si $x = \sum_{i\in I} x_i e_i \in E$ est orthogonal à tout élément de W, il est orthogonal à tous les e_i^*, $i \in I$. Par suite, pour tout $i \in I$, on a

$$0 = \langle x, e_i^* \rangle = x_i,$$

et ainsi $x = 0_E$.

Cela montre que la non dégénérescence de b ne suffit pas pour avoir une correspondance entre les sous-espaces de E et ceux de F. Pour en rajouter encore à la confusion ambiante, signalons que, dans le cas du crochet de dualité, on a une « moitié de correspondance », à savoir le résultat positif suivant.

2.13. Proposition. *Soit E un K-espace vectoriel. Alors, pour tout sous-espace V de E, on a ${}^\circ(V^\circ) = V$.*

En particulier :

(1) *l'application $V \longmapsto V^\circ$ est toujours injective ;*

(2) *l'application $W \longmapsto {}^\circ W$ est toujours surjective.*

Démonstration. Soit V un sous-espace de E, et soit V' un supplémentaire de V dans E. Par la remarque 2.6, on sait déjà que $V \subset {}^\circ(V^\circ)$. Soit maintenant $x \in {}^\circ(V^\circ)$. Écrivons $x = x_V + x_{V'}$, avec $x_V \in V$ et $x_{V'} \in V'$. Pour toute forme linéaire $\psi : V' \longrightarrow K$, on note $\widetilde{\psi} \in E^*$ l'unique forme linéaire sur E telle que

$$\widetilde{\psi}_{|V} = 0 \quad \text{et} \quad \widetilde{\psi}_{|V'} = \psi.$$

Par définition, $\widetilde{\psi} \in V^\circ$. Puisque $x \in {}^\circ(V^\circ)$, on a $\widetilde{\psi}(x) = 0$, soit

$$\widetilde{\psi}(x) = \widetilde{\psi}(x_V) + \widetilde{\psi}(x_{V'}) = 0_E + \psi(x_{V'}) = \psi(x_{V'}) = 0.$$

Autrement dit, on a

$$\psi(x_{V'}) = 0, \quad \text{pour tout } \psi \in (V')^*.$$

Le crochet de dualité sur V' étant non dégénéré par l'exemple 2.8, cela implique que $x_{V'} = 0_{V'}$, d'où $x \in V$. Ainsi, ${}^\circ(V^\circ) = V$. En particulier, tout

sous-espace V de E s'écrit $V = {}^\circ W$, où W est est un sous-espace de F, d'où la surjectivité de l'application $W \longmapsto {}^\circ W$. Si maintenant V_1 et V_2 sont deux sous-espaces de E vérifiant $V_1^\circ = V_2^\circ$, on obtient ${}^\circ(V_1^\circ) = {}^\circ(V_2^\circ)$, soit encore $V_1 = V_2$. Cela achève la démonstration. □

2.14. Remarque. Si E est de dimension infinie, l'application $V \longmapsto V^\circ$ n'est pas surjective.

En effet, supposons que E soit de dimension infinie, de base $\mathbf{e} = (e_i)_{i\in I}$, et soit

$$W = \mathrm{Vect}_K(e_i^*,\ i \in I).$$

Rappelons que $W \neq E^*$, et que $({}^\circ W)^\circ = E^*$, par l'exemple 2.12 (2).

Si l'on avait $W = V^\circ$ pour un certain sous-espace V de E, la proposition précédente impliquerait l'égalité $V = {}^\circ W$. Mais alors, on aurait

$$W = V^\circ = ({}^\circ W)^\circ = E^*,$$

d'où une contradiction.

Ce même exemple montre aussi que l'application $W \longmapsto {}^\circ W$ n'est pas injective en dimension infinie.

En effet, on a ${}^\circ(E^*) = \{0_E\}$, puisque le seul vecteur de $x \in E$ annulant toute forme linéaire $\varphi \in E^*$ est le vecteur nul par l'exemple 2.8. Mais on sait aussi que ${}^\circ W = \{0_E\}$ (cf. exemple 2.12 (2) ci-dessus). Par conséquent, on a ${}^\circ W = {}^\circ(E^*)$, tandis que $W \neq E^*$.

Dans la suite, nous allons expliquer le pourquoi de ces différences entre dimension finie et infinie. Avant cela, nous allons faire une pause et introduire brièvement la notion d'espace vectoriel quotient.

3. Interlude : espaces vectoriels quotient

Soit E un K-espace vectoriel, et soit V un sous-espace de E. Alors, E étant un groupe abélien, on peut former le groupe quotient E/V.

Si maintenant $\omega \in E/V$ est une classe d'équivalence, et si $x \in E$ en est un représentant, pour tout $\lambda \in K$, on pose

$$\lambda\cdot\omega = \overline{\lambda x}.$$

Si $x' \in E$ est un autre représentant de ω, alors $\overline{x} = \overline{x'}$, soit $x' - x \in V$. Mais alors, on a

$$\lambda x' - \lambda x = \lambda(x' - x) \in V,$$

car V est un sous-espace de E. Autrement dit, la classe d'équivalence $\overline{\lambda x}$ ne dépend que de $\overline{x}$, et non pas du choix d'un représentant $x \in E$.

Cela permet de définir une loi externe $K \times E/V \longrightarrow E/V$ telle que

$$\lambda \cdot \overline{x} = \overline{\lambda x}, \quad \text{pour tout } x \in V \text{ et tout } \lambda \in K.$$

De même que pour les groupes et les anneaux quotients, on a alors le résultat suivant, dont la démonstration découle simplement des définitions.

3.1. Théorème. *Soit E un K-espace vectoriel, et soit V un sous-espace vectoriel de E. Le groupe quotient E/V, muni de la loi externe*

$$\begin{aligned} K \times E/V &\longrightarrow E/V \\ (\lambda, \overline{x}) &\longmapsto \lambda \cdot \overline{x} = \overline{\lambda x}, \end{aligned}$$

est un K-espace vectoriel.

De plus, la projection canonique $\pi : E \longrightarrow E/V$ est linéaire.

Tous les résultats sur les structures quotient rencontrées jusqu'à présent sont encore vrais dans ce cadre.

3.2. Proposition. *Soit E un K-espace vectoriel, et soit V un sous-espace vectoriel de E. Alors, la projection canonique $\pi : E \longrightarrow E/V$ induit une correspondance bijective entre l'ensemble des sous-espaces vectoriels V' de E contenant V et l'ensemble des sous-espaces vectoriels de E/V.*

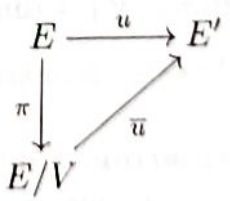
Plus précisément, si V' est un sous-espace vectoriel de E contenant F, alors

$$V'/V = \pi(V') = \{\overline{v'} \mid v' \in V'\}$$

est un sous-espace vectoriel de E/V. Réciproquement, si W est un sous-espace vectoriel de E/V, alors $\pi^{-1}(W)$ est un sous-espace vectoriel de E contenant V', et ces deux constructions sont inverses l'une de l'autre.

3.3. Théorème (de factorisation). *Soit E un K-espace vectoriel, et soit V un sous-espace vectoriel de E. Soit $u : E \longrightarrow E'$ une application linéaire telle que $V \subset \mathrm{Ker}(u)$. Alors, il existe une unique application linéaire*

$$\overline{u} : E/V \longrightarrow E'$$



telle que $u = \overline{u} \circ \pi$.
Cette application linéaire est définie par

$$\overline{u}(\overline{x}) = u(x), \quad \text{pour tout } \overline{x} \in E/V.$$

De plus, pour tout K-espace vectoriel E', il y a un isomorphisme d'espaces vectoriels entre $\mathscr{L}(E/V, E')$ et $\{u \in \mathscr{L}(E, E') \mid V \subset \mathrm{Ker}(u)\}$.

3.4. Théorème (premier théorème d'isomorphisme)

Soit $u \in \mathscr{L}(E, E')$ une application linéaire.

Alors, l'application linéaire $\overline{u} : E/\mathrm{Ker}(u) \longrightarrow E'$ est injective, et induit un isomorphisme d'espaces vectoriels

$$E/\mathrm{Ker}(u) \simeq \mathrm{Im}(u).$$

Ces trois résultats ont déjà été établis pour les groupes quotients (cf. proposition IX-1.6, théorème IX-2.1 et théorème IX-2.3). Il suffit donc de vérifier que tous les sous-groupes (resp. morphismes de groupes) considérés sont des sous-espaces vectoriels (resp. applications linéaires), ce qui est laissé au lecteur.

Nous allons maintenant introduire la notion de codimension d'un espace vectoriel.

3.5. Définition. Soit E un K-espace vectoriel, et soit V un sous-espace de E. La *codimension* de V est la dimension de l'espace E/V.

On la note $\mathrm{codim}_K(V)$.

On finit en calculant la dimension d'un espace quotient.

3.6. Proposition. *Soit E un K-espace vectoriel, et soit V un sous-espace vectoriel. Si V' est un supplémentaire de V, on a $E/V \simeq V'$. En particulier, si E est de dimension finie, V est de codimension finie, et l'on a*

$$\mathrm{codim}_K(V) = \dim_K(E) - \dim_K(V).$$

Démonstration. Soit V un sous-espace de V, et soit V' un supplémentaire de V dans E. On a donc $E = V \oplus V'$. Alors, la projection $\pi_{V'}$ sur V' parallèlement à V est linéaire, surjective, de noyau V. Le théorème de factorisation nous donne alors un isomorphisme $E/V \simeq V'$.

Si maintenant E est de dimension finie, il en est de même de V', donc de E/V, et l'on a alors

$$\mathrm{codim}_K(V) = \dim_K(E/V) = \dim_K(V') = \dim_K(E) - \dim_K(V).$$

Cela achève la démonstration. □

3.7. Remarque. Cette proposition permet de donner une autre caractérisation d'un hyperplan : c'est un sous-espace de codimension 1.

En effet, si H est un hyperplan de E, alors il existe une droite D telle que $E = H \oplus D$. Par la proposition précédente, on a $E/H \simeq D$, et donc $\mathrm{codim}_K(H) = 1$. Inversement, si V est un sous-espace de codimension 1, et

si V' est un supplémentaire de V dans E, on a $E/V \simeq V'$. Mais alors, par hypothèse, on a $\dim_K(V') = 1$, et V' est donc une droite de E. Ainsi, V est un hyperplan.

4. Dualité entre sous-espaces

Soit $b : E \times F \longrightarrow K$ une forme bilinéaire.

Dans cette section, nous allons finalement examiner de plus près le problème de l'existence de la correspondance souhaitée entre les sous-espaces de E et ceux de F.

On commence par donner une description de ${}^{\perp}(V^{\perp})$ en termes d'orthogonalité par rapport au crochet de dualité.

4.1. Lemme. *Soit $b : E \times F \longrightarrow K$ une forme bilinéaire. Alors, pour tout sous-espace V de E, on a*

$$ {}^{\perp}(V^{\perp}) = {}^{\circ}\big(V^{\circ} \cap \mathrm{Im}(b_*)\big). $$

Démonstration. Soit V un sous-espace de E. On a

$$\begin{aligned}
{}^{\perp}(V^{\perp}) &= \{x \in E \mid b(x,y) = 0, \text{ pour tout } y \in V^{\perp}\} \\
&= \{x \in E \mid b_*(y)(x) = 0, \text{ pour tout } y \in V^{\perp}\} \\
&= \{x \in E \mid b_*(y)(x) = 0, \text{ pour tout } y \in F \text{ tel que } \big(b_*(y)\big)_{|V} = 0\} \\
&= \{x \in E \mid b_*(y)(x) = 0, \text{ pour tout } y \in F \text{ tel que } b_*(y) \in V^{\circ}\} \\
&= \{x \in E \mid \varphi(x) = 0, \text{ pour tout } \varphi \in V^{\circ} \cap \mathrm{Im}(b_*)\} \\
&= {}^{\circ}\big(V^{\circ} \cap \mathrm{Im}(b_*)\big).
\end{aligned}$$

□

Le lemme suivant établit l'existence de quelques applications linéaires dont nous aurons besoin par la suite.

4.2. Lemme. *Soit $b : E \times F \longrightarrow K$ une application bilinéaire.*

(1) *Pour tout sous-espace V de E, l'application*

$$\begin{aligned} F &\longrightarrow V^* \\ y &\longmapsto b_*(y)_{|V} \end{aligned}$$

induit une application linéaire injective

$$ \rho_V : F/V^{\perp} \hookrightarrow V^*. $$

De plus, pour tout $y \in V^{\perp}$, la forme linéaire $b_*(y)$ induit une forme

linéaire $\overline{b_*(y)} \in (E/V)^*$, *et l'application*

$$\begin{aligned} \theta_V : V^\perp &\longrightarrow (E/V)^* \\ y &\longmapsto \overline{b_*(y)} \end{aligned}$$

est linéaire.

Si b_* *est injective (resp. surjective, bijective), il en est de même de* θ_V.

(2) *Pour tout sous-espace* W *de* F, *l'application*

$$\begin{aligned} E &\longrightarrow W^* \\ x &\longmapsto {}_*b(x)_{|W} \end{aligned}$$

induit une application linéaire injective

$$\rho_W : E/{}^\perp W \hookrightarrow W^*.$$

De plus, pour tout $x \in {}^\perp W$, *la forme linéaire* ${}_*b(x)$ *induit une forme linéaire* $\overline{{}_*b(x)} \in (F/W)^*$, *et l'application*

$$\begin{aligned} \theta_W : W^\perp &\longrightarrow (F/W)^* \\ x &\longmapsto \overline{{}_*b(x)} \end{aligned}$$

est linéaire.

Si ${}_*b$ *est injective (resp. surjective, bijective), il en est de même de* θ_W.

Démonstration. Il suffit de démontrer le premier point, le second se démontrant de manière similaire. Soit V un sous-espace de E. Par définition même, le noyau de l'application

$$\begin{aligned} F &\longrightarrow V^* \\ y &\longmapsto b_*(y)_{|V} \end{aligned}$$

est $V^\perp$. Par le premier théorème d'isomorphisme, elle induit donc une application linéaire injective

$$\rho_V : F/V^\perp \hookrightarrow V^*.$$

Soit maintenant $y \in V^\perp$. Alors, la forme linéaire $b_*(y) : E \longrightarrow K$ s'annule sur V, et induit donc une forme linéaire $\overline{b_*(y)} : E/V \longrightarrow K$ par le théorème de factorisation.

La linéarité de b_* implique facilement la linéarité de l'application linéaire

$$\begin{aligned} \theta_V : V^\perp &\longrightarrow (E/V)^* \\ y &\longmapsto \overline{b_*(y)}. \end{aligned}$$

Supposons que b_* soit injective, et montrons que l'application θ_V est injective. Soit $y \in V^\perp$ tel que $\overline{b_*(y)}$ soit la forme linéaire nulle. Pour tout $x \in E$,

on a alors $$b_*(y)(x) = \overline{b_*(y)}(\overline{x}) = 0.$$

Ainsi, $b_*(y)$ est nulle, et comme b_* est injective, on en déduit que $y = 0$. Par conséquent, θ_V est injective.

Supposons que b_* soit surjective et soit $\psi \in (E/V)^*$. Alors, $\varphi = \psi \circ \pi$ est une forme linéaire sur E s'annulant sur V. Par hypothèse, il existe $y \in F$ tel que $b_*(y) = \varphi$. Puisque φ s'annule sur V, on a donc $y \in V^\perp$. Mais alors, pour tout $x \in E$, on a

$$\overline{b_*(y)}(\overline{x}) = b_*(y)(x) = \varphi(x) = \psi(\overline{x}).$$

Ainsi, $\theta_V(y) = \psi$, et θ_V est surjective.

Il découle immédiatement de ce qui précède que si b_* est bijective, θ_V est un isomorphisme, d'où le résultat. □

Notation. Si E est un K-espace vectoriel et $k \geqslant 0$, on note $\mathcal{V}_k(E)$ l'ensemble des sous-espaces de E de dimension k, et $\mathcal{V}^k(E)$ l'ensemble des sous-espaces de E de codimension k.

Nous pouvons maintenant énoncer et démontrer le résultat principal de cette section.

4.3. Théorème. *Soit $b : E \times F \longrightarrow K$ une application bilinéaire.*

(1) *On a ${}^\perp(V^\perp) = V$ pour tout sous-espace V de E si, et seulement si, b_* est surjective.*
Dans ce cas, ${}_*b$ est injective.
Si de plus b_* est bijective, alors on a une correspondance bijective

$$\begin{array}{ccc} \mathcal{V}^k(E) & \longleftrightarrow & \mathcal{V}_k(F) \\ V & \longmapsto & V^\perp \\ {}^\perp W & \longleftarrow & W \end{array}$$

entre les sous-espaces de codimension k de E et les sous-espaces de dimension k de F.

(2) *On a $({}^\perp W)^\perp = W$ pour tout sous-espace W de F si, et seulement si, ${}_*b$ est surjective.*
Dans ce cas, b_* est injective.
Si de plus ${}_*b$ est bijective, alors on a une correspondance bijective

$$\begin{array}{ccc} \mathcal{V}_k(E) & \longleftrightarrow & \mathcal{V}^k(F) \\ V & \longmapsto & V^\perp \\ {}^\perp W & \longleftarrow & W \end{array}$$

entre les sous-espaces de dimension k de E et les sous-espaces de codimension k de F.

Démonstration. Encore une fois, il suffit de démontrer le premier point.

Supposons tout d'abord que b_* soit surjective. Alors, pour tout $V \in E$, le lemme 4.1 entraîne

$$^\perp(V^\perp) = {}^\circ(V^\circ) = V,$$

la dernière égalité découlant de la proposition 2.13. De plus, ${}_*b$ est injective par le lemme[7] 2.11 (1).

Supposons maintenant que b_* ne soit pas surjective, et soit $\varphi \in E^* \setminus \mathrm{Im}(b_*)$. Posons $V = \mathrm{Ker}(\varphi)$. Notons que $V \neq E$, puisque φ est en particulier non nulle. Par le lemme III-5.8, on a $V^\circ = K\varphi$. Comme $\varphi \notin \mathrm{Im}(b_*)$, on a[8]

$$V^\circ \cap \mathrm{Im}(b_*) = \{0_{E^*}\}.$$

Mais alors, par le lemme 4.1, on a

$$^\perp(V^\perp) = {}^\circ\{0_{E^*}\} = E \neq V.$$

Supposons que b_* soit bijective. Par le lemme 4.2 (1), l'application

$$\begin{aligned} \theta_V : V^\perp &\longrightarrow (E/V)^* \\ y &\longmapsto \overline{b_*(y)} \end{aligned}$$

est un isomorphisme.

Soit V un sous-espace de E de codimension k. Puisque $V^\perp \simeq (E/V)^*$, $V^\perp$ est de dimension finie, et l'on a

$$\dim_K(V^\perp) = \dim_K\left((E/V)^*\right) = \dim_K(E/V) = \mathrm{codim}_K(V) = k.$$

De plus, b_* étant bijective, donc surjective, on a $^\perp(V^\perp) = V$.

Pour conclure, il reste donc à montrer que, pour tout sous-espace W de F de dimension k, $^\perp W$ est de codimension k et $(^\perp W)^\perp = W$.

Par le lemme 4.2 (2), on a une application injective

$$\rho_W : E/{}^\perp W \hookrightarrow W^*.$$

L'espace $E/{}^\perp W$ est donc de dimension finie, et l'on a

$$\mathrm{codim}_K({}^\perp W) \leqslant k.$$

Par ailleurs, le point précédent montre que l'on a

$$\dim_K\left((^\perp W)^\perp\right) = \mathrm{codim}_K({}^\perp W).$$

On en déduit que $\dim_K\left((^\perp W)^\perp\right) \leqslant k$. Mais, on a aussi $W \subset (^\perp W)^\perp$, d'où l'inégalité $\dim_K\left((^\perp W)^\perp\right) \geqslant k$. Finalement, on obtient

$$\dim_K\left((^\perp W)^\perp\right) = k = \dim_K(W).$$

7. On peut aussi appliquer l'égalité précédente à $V = \{0_E\}$.
8. En effet, si $\mathrm{Im}(b_*)$ contient un multiple scalaire de φ non nul, il contient φ.

Comme $W \subset ({}^{\perp}W)^{\perp}$ et que ces deux sous-espaces sont de dimension finie, on obtient l'égalité souhaitée. Cela achève la démonstration. □

4.4. Exemple. Soit E un K-espace vectoriel. Regardons ce que donne le théorème pour le crochet de dualité $\langle\ ,\ \rangle_E$.

L'application $(\langle\ ,\ \rangle_E)_*$ est Id_{E^*}, qui est bijective. Ainsi, on retrouve le résultat de la proposition 2.13, à savoir

$$^{\circ}(V^{\circ}) = V,$$

pour tout sous-espace V de E. On a également une correspondance bijective

$$\begin{array}{ccc} \mathcal{V}^k(E) & \longleftrightarrow & \mathcal{V}_k(E^*) \\ V & \longmapsto & V^{\circ} \\ {}^{\circ}W & \longleftarrow & W \end{array}$$

entre les sous-espaces de E de codimension k et les sous-espaces de E^* de dimension k, et cela pour tout $k \geqslant 0$.

L'égalité $({}^{\circ}W)^{\circ} = W$ pour tout sous-espace de E^* de dimension k se retraduit de la manière suivante : pour toute famille $(\varphi_1, \ldots, \varphi_k)$ de formes linéaires sur E linéairement indépendantes, et pour tout $\varphi \in E^*$, on a

$$\bigcap_{i=1}^{k} \mathrm{Ker}(\varphi_i) \subset \mathrm{Ker}(\varphi) \iff \varphi \in \mathrm{Vect}_K(\varphi_1, \ldots, \varphi_k).$$

En effet, ce résultat est exactement l'égalité

$$({}^{\circ}W)^{\circ} = W,$$

où $W = \mathrm{Vect}_K(\varphi_1, \ldots, \varphi_k)$, une fois que l'on a remarqué que l'on a

$$\bigcap_{\varphi \in W} \mathrm{Ker}(\varphi) = \bigcap_{i=1}^{k} \mathrm{Ker}(\varphi_i),$$

ce qui est un exercice facile laissé au lecteur.

Notons enfin que cette équivalence peut être démontrée directement par récurrence (cf. exercice 15).

Rappelons maintenant que ${}_*\langle\ ,\ \rangle_E$ est l'application $\chi : E \longrightarrow E^{**}$ qui à un vecteur $x \in E$ associe l'évaluation en x, i.e.

$$\chi(x)(\varphi) = \varphi(x), \quad \text{pour tout } \varphi \in E^*.$$

On sait que l'égalité $({}^{\circ}W)^{\circ} = W$ est vraie en dimension finie d'après ce qui précède, et fausse en dimension infinie (cf. exemple 2.12 (2)). Cela s'explique par le fait que $\chi : E \hookrightarrow E^{**}$ est surjective si, et seulement si, E est de dimension finie (cf. exemple 2.10).

On continue en donnant une condition nécessaire et suffisante pour avoir notre correspondance entre **tous** les sous-espaces de E et F.

4.5. Théorème. *Soit $b : E \times F \longrightarrow K$ une forme bilinéaire. Les propriétés suivantes sont équivalentes :*

(1) *pour tout sous-espace V de E, et pour tout sous-espace W de F, on a*

$$^{\perp}(V^{\perp}) = V \quad \text{et} \quad (^{\perp}W)^{\perp} = W\,;$$

(2) *la forme bilinéaire b est non dégénérée, et E et F sont de dimension finie.*

Dans ce cas, $\dim_K(E) = \dim_K(F)$, et si l'on note n leur dimension, pour tout $k \in [\![0, n]\!]$, on a une correspondance bijective

$$\begin{array}{ccc} \mathcal{V}_{n-k}(E) & \longleftrightarrow & \mathcal{V}_k(F) \\ V & \longmapsto & V^{\perp} \\ {}^{\perp}W & \longleftarrow & W \end{array}$$

entre les sous-espaces de dimension $n-k$ de E et les sous-espaces de dimension k de F.

Démonstration

(1) $\Longrightarrow$ (2). Puisque l'on a l'égalité $^{\perp}(V^{\perp}) = V$ pour tout sous-espace V de E, le théorème 4.3 implique que b_* est surjective, et dans ce cas ${}_*b$ est injective. De même, $(^{\perp}W)^{\perp} = W$ pour tout sous-espace W de F, ${}_*b$ est surjective, et b_* est injective. Ainsi, les applications b_* et ${}_*b$ sont bijectives. Le lemme 2.11 (3) montre alors que b est non dégénérée, et que E et F sont de dimension finie.

(2) $\Longrightarrow$ (1). Par le lemme 2.11 (3), b_* et ${}_*b$ sont bijectives. Il suffit alors d'appliquer le théorème 4.3, en remarquant qu'un sous-espace de E est de codimension k si, et seulement si, il est de dimension $n-k$ par la proposition 3.6. □

Ce résultat est en quelque sorte un résultat négatif. Il dit en substance que si l'on veut établir une correspondance bijective entre les sous-espaces de E et ceux de F via une forme bilinéaire $b : E \times F \longrightarrow K$, on est forcé de se placer en dimension finie.

On finit cette section en généralisant la décomposition $E = F \oplus F^{\perp}$ bien connue dans le cas euclidien.

4.6. Proposition. *Soit $b : E \times F \longrightarrow K$ une forme bilinéaire. Soient V et W des sous-espaces de E et F respectivement.*

(1) *Si ${}_*(b_{|V\times W}) : V \longrightarrow W^*$ est bijective, alors $E = V \oplus {}^\perp W$;*

(2) *si $(b_{|V\times W})_* : W \longrightarrow V^*$ est bijective, alors $F = W \oplus V^\perp$.*

En particulier, ces deux décompositions sont valables pour tous sous-espaces de dimension finie V et W tels que $b_{|V\times W} : V \times W \longrightarrow K$ soit non dégénérée.

Démonstration. Il suffit de démontrer (1), le point (2) se démontrant de même. Soit $x \in E$. Par surjectivité de ${}_*(b_{|V\times W})$, il existe un vecteur $v \in V$ tel que ${}_*(b_{|V\times W})(v) = \big({}_*b(x)\big)_{|W}$. Cela se récrit

$$b(v, w) = b(x, w), \qquad \text{pour tout } w \in W.$$

On a donc $b(x - v, w) = 0$ pour tout $w \in W$, et ainsi $x - v \in {}^\perp W$. Puisque l'on a $x = v + (x - v)$, on obtient

$$E = V + {}^\perp W.$$

Soit maintenant $v \in V \cap {}^\perp W$. Alors, pour tout $w \in W$, on a $b(v, w) = 0$, et donc $v \in \mathrm{Ker}\big({}_*(b_{|V\times W})\big)$. Par hypothèse, $v = 0$. Cela démontre le résultat attendu. La dernière partie découle du lemme 2.11 (3). Cela achève la démonstration. □

4.7. Remarque. Si $E = F$ et $V = W$, les décompositions ci-dessus, lorsqu'elles sont valables, sont « orthogonales », dans le sens où les deux composantes d'un vecteur $x \in E$ sont orthogonales par rapport à b.

4.8. Exemple. Soit $b : E \times E \longrightarrow K$ une forme bilinéaire symétrique et soit V un sous-espace de E de dimension finie. Alors, on a $E = V \oplus V^\perp$ si, et seulement si, $b_{|V\times V} : V \times V \longrightarrow K$ est non dégénérée.

En effet, si $b_{|V\times V} : V \times V \longrightarrow K$ est non dégénérée, il suffit d'appliquer le théorème précédent. Inversement, supposons que $E = V \oplus V^\perp$. En particulier, on a $V \cap V^\perp = \{0_E\}$. Soit $x \in V$ tel que $b(x, y) = 0$, pour tout $y \in V$. On a donc $x \in V^\perp$, d'où $x = 0_E$ par hypothèse. Ainsi, b est non dégénérée à gauche, et par symétrie, b est aussi non dégénérée à droite, d'où le résultat.

En particulier, si E est un espace vectoriel réel, muni d'un produit scalaire

$$\langle\ ,\ \rangle : E \times E \longrightarrow \mathbb{R},$$

pour tout sous-espace V de E, la restriction de $\langle\ ,\ \rangle$ à $V \times V$ est un produit scalaire sur V, donc une forme bilinéaire non dégénérée. En particulier,

pour tout sous-espace V de dimension finie, on a

$$E = V \oplus V^{\perp}.$$

Remarquons que tout cela devient faux si V est de dimension infinie. En effet, choisissons E et V comme dans l'exemple 2.12 (1). Alors, $E \neq V$, mais $V^{\perp} = \{0_E\}$. En particulier, on ne peut avoir $E = V \oplus V^{\perp}$.

Nous allons maintenant établir une correspondance entre les bases de E et celles de F, dans le cas où E et F sont de dimension finie.

5. Bases duales

Dans ce qui suit, E et F seront des K-espaces vectoriels non nuls de dimension finie, et $b : E \times F \longrightarrow K$ sera une forme bilinéaire non dégénérée. Le but de cette courte section est d'établir une correspondance entre les bases de E et celles de F.

On commence par étudier à quelles conditions b est non dégénérée. Pour cela, on introduit la notion de matrice représentative d'une forme bilinéaire.

5.1. Définition. Soient E, F deux espaces vectoriels de dimension finie, et soit $b : E \times F \longrightarrow K$ une forme bilinéaire (pas nécessairement non dégénérée !). Soient $\mathbf{e} = (e_1, \ldots, e_n)$ et $\mathbf{f} = (f_1, \ldots, f_m)$ des bases de E et F respectivement. La *matrice représentative* de b dans les bases $\mathbf{e}$ et $\mathbf{f}$ est la matrice

$$\mathrm{Mat}(b\,; \mathbf{e}, \mathbf{f}) = \big(b(e_i, f_j)\big)_{i,j}.$$

5.2. Lemme. *Soient E, F deux espaces vectoriels de dimension finie, et soit $b : E \times F \longrightarrow K$ une forme bilinéaire.*

Soient $\mathbf{e} = (e_1, \ldots, e_n)$ et $\mathbf{f} = (f_1, \ldots, f_m)$ des bases de E et F respectivement, et soit $B = \mathrm{Mat}(b\,; \mathbf{e}, \mathbf{f})$. Alors, on a

$$b(x, y) = [x]_{\mathbf{e}}^{t} B [y]_{\mathbf{f}}.$$

Démonstration. Écrivons $x = \sum_{i=1}^{n} x_i e_i$ et $y = \sum_{j=1}^{m} y_j f_j$. Par bilinéarité, on a

$$b(x, y) = b\Big(\sum_{i=1}^{n} x_i e_i, \sum_{j=1}^{m} y_j f_j\Big) = \sum_{i,j} x_i b(e_i, f_j) y_j = [x]_{\mathbf{e}}^{t} B [y]_{\mathbf{f}}.$$

Les détails sont laissés au lecteur. □

5.3. Proposition. *Soient E, F deux espaces vectoriels de dimension finie, et soit $b : E \times F \longrightarrow K$ une forme bilinéaire.*

Soient $\mathbf{e} = (e_1, \ldots, e_n)$ et $\mathbf{f} = (f_1, \ldots, f_m)$ des bases de E et F respectivement. Alors, on a

$$\mathrm{Mat}(b\,;\mathbf{e},\mathbf{f}) = \mathrm{Mat}(b_*\,;\mathbf{f},\mathbf{e}^*) = \mathrm{Mat}({}_*b\,;\mathbf{e},\mathbf{f}^*)^t.$$

En particulier, la forme bilinéaire b est non dégénérée si, et seulement si, sa matrice représentative dans des bases arbitraires de E et F est inversible.

Démonstration. Pour tout $j \in [\![1, m]\!]$, on a

$$b_*(f_j) = b(\_, f_j) = \sum_{i=1}^{n} b(e_i, f_j) e_i^*,$$

d'après la proposition III-5.4. De même pour tout $k \in [\![1, n]\!]$, on a

$${}_*b(e_k) = b(e_k, \_\,) = \sum_{i=1}^{m} b(e_k, f_i) f_i^*,$$

d'où la première assertion.

Si b est non dégénérée, on a vu que b_* était un isomorphisme. Sa matrice représentative est donc inversible, et le point précédent montre alors que la matrice $\mathrm{Mat}(b\,;\mathbf{e},\mathbf{f})$ est inversible pour toutes bases $\mathbf{e}$ et $\mathbf{f}$. Inversement, supposons que $\mathrm{Mat}(b\,;\mathbf{e},\mathbf{f})$ soit inversible. Les égalités précédentes montrent alors que les matrices $\mathrm{Mat}(b_*;\mathbf{f},\mathbf{e}^*)$ et $\mathrm{Mat}({}_*b;\mathbf{e},\mathbf{f}^*)$ sont inversibles, et donc que ${}_*b$ et b_* sont des isomorphismes.
En particulier, la forme bilinéaire $b : E \times F \longrightarrow K$ est non dégénérée. □

Nous allons maintenant définir la notion de bases duales par rapport à b.

5.4. Définition. Soient E et F deux espaces vectoriels de même dimension finie $n \geqslant 1$ sur K, et soit $b : E \times F \longrightarrow K$ une forme bilinéaire.

Deux bases $\mathbf{e} = (e_1, \ldots, e_n)$ et $\mathbf{f} = (f_1, \ldots, f_n)$ de E et F respectivement sont dites *duales* par rapport à $b : E \times F \longrightarrow K$ si l'on a

$$b(e_i, f_j) = \delta_{ij}, \qquad \text{pour tous } i, j \in [\![1, n]\!].$$

Autrement dit, $\mathbf{e}$ et $\mathbf{f}$ sont duales par rapport à b si, et seulement si, on a

$$\mathrm{Mat}(b\,;\mathbf{e},\mathbf{f}) = I_n \in \mathrm{M}_n(K).$$

5.5. Remarque. La proposition 5.3 implique en particulier que s'il existe deux bases $\mathbf{e}$ et $\mathbf{f}$ duales par rapport à b, alors b est nécessairement non dégénérée.

5.6. Exemple. Soit E un K-espace vectoriel de dimension finie, et soit $\mathbf{e}$ une base de E. Alors, les bases $\mathbf{e}$ et $\mathbf{e}^*$ sont duales par rapport au crochet de dualité.

Le lemme suivant explique l'intérêt de cette nouvelle notion.

5.7. Lemme. *Si* $\mathbf{e}$ *et* $\mathbf{f}$ *sont deux bases duales par rapport à* b, *alors pour tous* $x \in E$, $y \in F$, *on a*

$$x = b(x, f_1)e_1 + \cdots + b(x, f_n)e_n$$

et

$$y = b(e_1, y)f_1 + \cdots + b(e_n, y)f_n.$$

Démonstration. Écrivons $x = \sum_{i=1}^{n} x_i e_i$. Pour tout $j \in [\![1, n]\!]$, on a

$$b(x, f_j) = \sum_{i=1}^{n} x_i b(e_i, f_j) = x_j.$$

L'autre égalité se montre de la même manière. □

5.8. Exemples

(1) Soit E un K-espace vectoriel de dimension finie, et soit $\mathbf{e}$ une base de E. En appliquant ce résultat au crochet de dualité avec $\mathbf{f} = \mathbf{e}^*$, on retrouve les formules énoncées dans la proposition III-5.4.

(2) Soit $(E, \langle\ ,\ \rangle)$ un espace euclidien, et soit $\mathbf{e} = (e_1, \ldots, e_n)$ une base de E. Alors, $\mathbf{e}$ est orthonormée si, et seulement si, les bases $\mathbf{e}$ et $\mathbf{e}^*$ sont duales par rapport au produit scalaire. Si $\mathbf{e}$ est orthonormée, le lemme précédent redonne alors l'égalité classique

$$x = \sum_{j=1}^{n} \langle x, e_j \rangle e_j, \qquad \text{pour tout } x \in E.$$

On va maintenant montrer le résultat principal de cette section.

5.9. Théorème. *Soit* $b : E \times F \longrightarrow K$ *une forme bilinéaire non dégénérée, où* E *et* F *sont de même dimension finie. Alors, pour toute base* $\mathbf{e}$ *de* E, *il existe une unique base* $\mathbf{e}^*$ *de* F *telle que les bases* $\mathbf{e}$ *et* $\mathbf{e}^*$ *soient duales par rapport à* b. *De même, pour toute base* $\mathbf{f}$ *de* F, *il existe une unique base* ${}^*\mathbf{f}$ *de* E *telle que les bases* ${}^*\mathbf{f}$ *et* $\mathbf{f}$ *soient duales par rapport à* b. *De plus, on a*

$${}^*(\mathbf{e}^*) = \mathbf{e} \quad \text{et} \quad ({}^*\mathbf{f})^* = \mathbf{f}.$$

Démonstration. Soit $\mathbf{e}$ une base de E et soit $\mathbf{f}$ une base quelconque de F. Puisque b est non dégénérée, la matrice $B = \mathrm{Mat}(b\,; \mathbf{e}, \mathbf{f})$ est inversible par la proposition 5.3.

Écrivons $B^{-1} = (m_{ij})$ et pour $j \in [\![1, n]\!]$, posons

$$e_j^* = \sum_{k=1}^{n} m_{kj} f_k.$$

Puisque B^{-1} est inversible, la famille $\mathbf{e}^*$ est une base de F. De plus, pour tous $i, j \in [\![1, n]\!]$, on a

$$b(e_i, e_j^*) = b\Big(e_i, \sum_{k=1}^{n} m_{kj} f_k\Big) = \sum_{k=1}^{n} b(e_i, f_k) m_{kj} = \delta_{ij},$$

la dernière égalité provenant du fait que $BB^{-1} = I_n$. Ainsi, $\mathbf{e}$ et $\mathbf{e}^*$ sont duales.

Supposons que $\mathbf{f^{(1)}}$ et $\mathbf{f^{(2)}}$ soient deux bases de F telles que $\mathbf{e}$ et $\mathbf{f^{(k)}}$ soient duales pour $k = 1, 2$. Alors, pour tous $i, j \in [\![1, n]\!]$, on a

$$b(e_i, f_j^{(1)} - f_j^{(2)}) = b(e_i, f_j^{(1)}) - b(e_i, f_j^{(2)}) = \delta_{ij} - \delta_{ij} = 0.$$

Puisque b est non dégénérée, on en déduit que $f_j^{(1)} - f_j^{(2)} = 0$ pour tout entier $j \in [\![1, n]\!]$ et donc $\mathbf{f^{(1)}} = \mathbf{f^{(2)}}$.

Soit maintenant $\mathbf{f}$ une base de F et soit $\mathbf{e}$ une base quelconque de E. Puisque b est non dégénérée, la matrice $B = \mathrm{Mat}(b\,; \mathbf{e}, \mathbf{f})$ est inversible par la proposition 5.3.

Écrivons $B^{-1} = (m_{ij})$ et pour $j \in [\![1, n]\!]$, posons cette fois

$$^*f_j = \sum_{k=1}^{n} m_{jk} e_k.$$

Puisque $(B^{-1})^t$ est inversible, la famille $^*\mathbf{f}$ est une base de E. De plus, pour tous $i, j \in [\![1, n]\!]$, on a

$$b(^*f_i, f_j) = b\Big(\sum_{k=1}^{n} m_{ik} e_k, f_j\Big) = \sum_{k=1}^{n} m_{ik} b(e_k, f_j) = \delta_{ij},$$

la dernière égalité provenant du fait que $B^{-1}B = I_n$. Ainsi, $^*\mathbf{f}$ et $\mathbf{f}$ sont duales. L'unicité d'une base duale à $\mathbf{f}$ se démontre comme précédemment.

Montrons le dernier point. Par définition, $^*(\mathbf{e}^*)$ est l'unique base de E telle que $^*(\mathbf{e}^*)$ et $\mathbf{e}^*$ soient duales. Mais par définition, $\mathbf{e}$ et $\mathbf{e}^*$ sont duales, d'où l'égalité désirée. La deuxième se démontre de manière analogue. □

5.10. Remarques

(1) La notation $\mathbf{e}^*$ est un peu dangereuse. Si b n'est pas le crochet de dualité, $\mathbf{e}^*$ est une base de F et non pas de E^*, et en particulier n'est **pas** la base des formes coordonnées !!!

(2) La démonstration précédente fournit un moyen de calculer explicitement les bases duales.

(i) Si $\mathbf{e}$ est une base de E, on choisit une base arbitraire $\mathbf{f}$ de F, et l'on pose $B = \mathrm{Mat}(b\,;\mathbf{e},\mathbf{f})$. Alors, pour tout entier $j \in [\![1,n]\!]$, $[e_j^*]_{\mathbf{f}}$ est la j-ième colonne de B^{-1}.

(ii) Si $\mathbf{f}$ est une base de F, on choisit une base arbitraire $\mathbf{e}$ de E, et l'on pose $B = \mathrm{Mat}(b\,;\mathbf{e},\mathbf{f})$. Alors, pour tout entier $j \in [\![1,n]\!]$, $[f_j^*]_{\mathbf{e}}$ est la j-ième colonne de $(B^{-1})^t$.

5.11. Définition. Si $\mathbf{e}$ est une base de E, la base $\mathbf{e}^*$ est appelée la *base duale* de $\mathbf{e}$ (par rapport à b). De même, si $\mathbf{f}$ est une base de F, la base ${}^*\mathbf{f}$ est appelée la *base duale* de $\mathbf{f}$ (par rapport à b).

5.12. Exemples

(1) Si $\mathbf{e}$ est une base de E, la base duale par rapport au crochet de dualité associée est la base des formes coordonnées $\mathbf{e}^*$, d'où la notation choisie.

(2) Soit $E = \mathbb{R}_n[X]$ l'espace vectoriel réel des polynômes de degré inférieur ou égal à n, et soit $\mathbf{e} = (1, X, \ldots, X^n)$. Pour tout entier $j \in [\![0,n]\!]$, posons

$$\begin{aligned} f_j : E &\longrightarrow \mathbb{R} \\ P &\longmapsto \frac{P^{(j)}(0)}{j!}. \end{aligned}$$

Alors, $\mathbf{f} = (f_0, \ldots, f_n)$ est la base duale de $\mathbf{e}$ par rapport au crochet de dualité.

(3) Soit $E = \mathbb{R}_n[X]$, et soient $a_0, \ldots, a_n \in \mathbb{R}$ des réels **distincts**. Pour tout entier $j \in [\![1,n]\!]$, soit

$$\begin{aligned} f_j : E &\longrightarrow \mathbb{R} \\ P &\longmapsto P(a_j). \end{aligned}$$

Alors, $\mathbf{f}$ est une base de E^*, et sa base duale est $\mathbf{e} = (L_0, \ldots, L_n)$, où L_i est défini par

$$L_i = \prod_{j \neq i} \frac{X - a_j}{a_i - a_j}.$$

Les vérifications sont laissées en exercice.

Le lecteur trouvera quelques applications des bases duales dans les exercices.

On continue ce chapitre en établissant une correspondance entre les endomorphismes de E et ceux de F.

6. Adjoint d'un endomorphisme

Dans ce chapitre, on généralise la notion d'adjonction, bien connue dans le cadre euclidien ou hermitien, au cas d'une forme bilinéaire quelconque.

On commence par un lemme.

6.1. Lemme. *Soient E, F, G des K-espaces vectoriels, soit $u \in \mathscr{L}(E,F)$ et soit $w \in \mathscr{L}(E,G)$. Alors, il existe $v \in \mathscr{L}(F,G)$ tel que $w = v \circ u$ si, et seulement si,* $\mathrm{Ker}(u) \subset \mathrm{Ker}(w)$.

Démonstration. Si $w = v \circ u$, il est clair que $\mathrm{Ker}(u) \subset \mathrm{Ker}(w)$. Inversement, supposons que $\mathrm{Ker}(u) \subset \mathrm{Ker}(w)$. Le théorème de factorisation montre alors que w induit par passage au quotient une application linéaire

$$\overline{w} : E/\mathrm{Ker}(u) \longrightarrow G.$$

Soit $\overline{u} : E/\mathrm{Ker}(u) \xrightarrow{\sim} \mathrm{Im}(u)$ l'isomorphisme canonique induit par u.

Choisissons un supplémentaire F' de $\mathrm{Im}(u)$ dans F, et soit $v : F \longrightarrow G$ l'unique application linéaire telle que $v_{|\mathrm{Im}(u)} = \overline{w} \circ \overline{u}^{-1}$ et $v_{|F'} = 0$. Alors, pour tout $x \in E$, on a

$$v\big(u(x)\big) = v\big(\overline{u}(\overline{x})\big) = \overline{w}(\overline{x}) = w(x),$$

d'où l'égalité $w = v \circ u$. Cela achève la démonstration. □

6.2. Remarque. Supposons que $F = G = K$. Puisque les endomorphismes de K sont exactement les applications de la forme

$$\begin{aligned} K &\longrightarrow K \\ x &\longmapsto \lambda x, \end{aligned}$$

on retrouve le fait que, pour tous φ, $\psi \in E^*$, on a

$$\mathrm{Ker}(\varphi) \subset \mathrm{Ker}(\psi) \iff \text{il existe } \lambda \in K \text{ tel que } \psi = \lambda\varphi.$$

Entrons maintenant dans le vif du sujet.

6.3. Théorème. *Soit $b : E \times F \longrightarrow K$ une forme bilinéaire, où E et F sont non nuls. Les propriétés suivantes sont équivalentes :*

(1) *l'application linéaire b_* est bijective ;*

(2) *pour tout $u \in \mathscr{L}(E)$, il existe un unique $u^* \in \mathscr{L}(F)$ tel que*

$$b\big(u(x), y\big) = b\big(x, u^*(y)\big), \quad \text{pour tous } x \in E,\ y \in F.$$

Dans ce cas, on a

$$\mathrm{Ker}(u^*) = \mathrm{Im}(u)^\perp \quad \textit{et} \quad \mathrm{Im}(u^*) = \mathrm{Ker}(u)^\perp.$$

En particulier, u^* est de rang fini si, et seulement si, u est de rang fini, et on a alors

$$\mathrm{rg}(u^*) = \mathrm{rg}(u).$$

Enfin, on a

(i) $\mathrm{Id}_E^* = \mathrm{Id}_F$;
(ii) *pour tous* $u_1, u_2 \in \mathscr{L}(E)$, *on a* $(u_1 \circ u_2)^* = u_2^* \circ u_1^*$;
(iii) *l'application*

$$\begin{aligned} \mathscr{L}(E) &\longrightarrow \mathscr{L}(F) \\ u &\longmapsto u^* \end{aligned}$$

est linéaire.

De même, les propriétés suivantes sont équivalentes :

(1) *l'application linéaire* ${}_*b$ *est bijective ;*

(2) *pour tout* $\rho \in \mathscr{L}(F)$, *il existe un unique* ${}^*\rho \in \mathscr{L}(E)$ *tel que*

$$b\big(x, \rho(y)\big) = b({}^*\rho(x), y), \qquad \textit{pour tous } x \in E,\ y \in F.$$

Dans ce cas, on a

$$\mathrm{Ker}({}^*\rho) = {}^\perp\mathrm{Im}(\rho) \quad \textit{et} \quad \mathrm{Im}({}^*\rho) = {}^\perp\mathrm{Ker}(\rho).$$

En particulier, ${}^*\rho$ est de rang fini si, et seulement si, ρ est de rang fini, et l'on a alors

$$\mathrm{rg}({}^*\rho) = \mathrm{rg}(\rho).$$

Enfin, on a

(i) ${}^*\mathrm{Id}_F = \mathrm{Id}_E$;
(ii) *pour tous* $\rho_1, \rho_2 \in \mathscr{L}(E)$, *on a* ${}^*(\rho_1 \circ \rho_2) = {}^*\rho_2 \circ {}^*\rho_1$;
(iii) *l'application*

$$\begin{aligned} \mathscr{L}(F) &\longrightarrow \mathscr{L}(E) \\ \rho &\longmapsto {}^*\rho \end{aligned}$$

est linéaire.

Démonstration. Il suffit de démontrer la première partie, l'autre se démontrant de manière similaire.

Supposons tout d'abord que pour tout endomorphisme $u \in \mathscr{L}(E)$, il existe un unique endomorphisme $u^* \in \mathscr{L}(F)$ vérifiant les conditions du théorème. Si b était identiquement nulle, alors l'endomorphisme nul et Id_F vérifieraient tous deux les conditions demandées, ce qui contredirait l'unicité[9]. Ainsi, il existe $x_0 \in E$ et $y_0 \in F$ tels que $b(x_0, y_0) \neq 0$.

9. On a bien $\mathrm{Id}_F \neq 0$, car F est non nul.

Soit $\varphi \in E^*$, et soit

$$\begin{aligned} u : E &\longrightarrow E \\ x &\longmapsto \frac{\varphi(x)}{b(x_0, y_0)} x_0. \end{aligned}$$

Considérons l'endomorphisme $u^* \in \mathscr{L}(E)$ associé. En particulier, pour tout vecteur $x \in E$, on a

$$\begin{aligned} b_*\big(u^*(y_0)\big)(x) &= b\big(x, u^*(y_0)\big) \\ &= b\big(u(x), y_0\big) \\ &= \frac{\varphi(x)}{b(x_0, y_0)} b(x_0, y_0) \\ &= \varphi(x). \end{aligned}$$

Ainsi, $\varphi = b_*\big(u^*(y_0)\big) \in \mathrm{Im}(b_*)$ et b_* est surjective. Montrons maintenant que l'application b_* est injective. Soit F' un supplémentaire de $\mathrm{Ker}(b_*)$ dans F, et soit $\pi \in \mathscr{L}(F)$ la projection sur $\mathrm{Ker}(b_*)$ parallèlement à F'. Par définition, pour tout $y \in F$, on a $\pi(y) \in \mathrm{Ker}(b_*)$, et ainsi

$$b\big(x, \pi(y)\big) = 0, \quad \text{pour tout } x \in E.$$

Autrement dit, si $u \in \mathscr{L}(E)$ est l'endomorphisme nul, $\pi \in \mathscr{L}(F)$ vérifie les conditions du théorème. Mais, il en est de même de l'endomorphisme nul de F. Par unicité, on obtient $\pi = 0$. En particulier, $\mathrm{Im}(\pi) = \mathrm{Ker}(b_*) = \{0\}$, et b_* est injective.

Inversement, supposons que l'application b_* soit bijective. Soit $u \in \mathscr{L}(E)$. Pour tout $y \in F$ fixé, l'application

$$\begin{aligned} \theta_y : E &\longrightarrow K \\ x &\longmapsto b\big(u(x), y\big) \end{aligned}$$

est une forme linéaire. De plus, l'application

$$\begin{aligned} \theta : F &\longrightarrow E^* \\ y &\longmapsto \theta_y \end{aligned}$$

est linéaire.

Posons $u^* = b_*^{-1} \circ \theta$. Alors, $u^* \in \mathscr{L}(F)$. De plus, pour tout vecteur $y \in F$, on a $b_*\big(u^*(y)\big) = \theta_y$. Ainsi, pour tout $x \in E$, on a

$$b\big(x, u^*(y)\big) = b\big(u(x), y\big),$$

d'où l'existence de u^*.

De plus, si u_1^*, $u_2^* \in \mathscr{L}(E)$ vérifient

$$b\big(u(x), y\big) = b\big(x, u_1^*(y)\big) = b\big(x, u_2^*(y)\big)$$

pour tous $x \in E$, $y \in F$, alors, pour tout $y \in F$, on a

$$b\big(x, u_1^*(y) - u_2^*(y)\big) = 0, \quad \text{pour tout } x \in E.$$

Ainsi, $u_1^*(y) - u_2^*(y) \in \mathrm{Ker}(b_*) = \{0\}$, d'où $u_1^*(y) = u_2^*(y)$. Cela étant vrai pour tout $y \in F$, on a $u_1^* = u_2^*$, d'où l'unicité.

Déterminons maintenant le noyau de u^*. Pour tout $y \in F$, on a

$$\begin{aligned} y \in \mathrm{Ker}(u^*) &\iff u^*(y) = 0 \\ &\iff b\big(x, u^*(y)\big) = 0, \text{ pour tout } x \in E \\ &\iff b\big(u(x), y\big) = 0, \text{ pour tout } x \in E \\ &\iff b(x', y) = 0, \text{ pour tout } x' \in \mathrm{Im}(u) \\ &\iff y \in \mathrm{Im}(u)^\perp, \end{aligned}$$

la deuxième équivalence découlant de l'injectivité de b_*. Par conséquent, on a bien $\mathrm{Ker}(u^*) = \mathrm{Im}(u)^\perp$.

Montrons que $\mathrm{Im}(u^*) = \mathrm{Ker}(u)^\perp$. Soit $y \in F$. Alors, pour tout $x \in \mathrm{Ker}(u)$, on a

$$b\big(x, u^*(y)\big) = b\big(u(x), y\big) = b(0_E, y) = 0.$$

Ainsi, $u^*(y) \in \mathrm{Ker}(u)^\perp$ pour tout $y \in F$, c'est-à-dire $\mathrm{Im}(u^*) \subset \mathrm{Ker}(u)^\perp$.

Soit $y \in \mathrm{Ker}(u)^\perp$. Alors, pour tout $x \in \mathrm{Ker}(u)$, on a $b_*(y)(x) = 0$. Autrement dit, $\mathrm{Ker}(u) \subset \mathrm{Ker}\big(b_*(y)\big)$. Par le lemme 6.1, il existe $\varphi \in E^*$ tel que $b_*(y) = \varphi \circ u$. Puisque l'application b_* est surjective, il existe $z \in F$ tel que $\varphi = b_*(z)$. On a donc $b_*(y) = b_*(z) \circ u$, et ainsi

$$b(x, y) = b\big(u(x), z\big), \quad \text{pour tout } x \in E.$$

Mais, alors on a

$$b(x, y) = b\big(x, u^*(z)\big), \quad \text{pour tout } x \in E.$$

Par injectivité de b_*, on a $y = u^*(z) \in \mathrm{Im}(u^*)$, d'où l'inclusion manquante.

Il reste à démontrer la dernière partie. Par le lemme 4.2, on a

$$V^\perp \simeq (E/V)^*, \quad \text{pour tout sous-espace } V \text{ de } E.$$

En particulier, on a

$$\mathrm{Im}(u^*) = \mathrm{Ker}(u)^\perp \simeq \big(E/\mathrm{Ker}(u)\big)^*.$$

Par le premier théorème d'isomorphisme, on a $E/\mathrm{Ker}(u) \simeq \mathrm{Im}(u)$. On en déduit alors un isomorphisme

$$\mathrm{Im}(u^*) \simeq \mathrm{Im}(u)^*.$$

En utilisant la proposition III-5.4, on en déduit les équivalences

$$\dim_K\big(\mathrm{Im}(u^*)\big) < +\infty \iff \dim_K(\mathrm{Im}(u)^*) < +\infty \iff \dim_K\big(\mathrm{Im}(u)\big) < \infty.$$

Ainsi, u^* est de rang fini si, et seulement si, u est de rang fini.

Dans ce cas, on a

$$\mathrm{rg}(u^*) = \dim_K\big(\mathrm{Im}(u^*)\big) = \dim_K\big(\mathrm{Im}(u)^*\big) = \dim_K\big(\mathrm{Im}(u)\big) = \mathrm{rg}(u).$$

Montrons enfin le dernier point. Pour tous $x \in E,\ y \in F$, on a

$$b\big(\mathrm{Id}_E(x), y\big) = b(x, y) = b\big(x, \mathrm{Id}_F(y)\big),$$

d'où $\mathrm{Id}_E^* = \mathrm{Id}_F$. Si $u_1,\ u_2 \in \mathscr{L}(E)$, alors pour tous $x \in E,\ y \in F$, on a

$$b\big(u_1\big(u_2(x)\big), y\big) = b\big(u_2(x), u_1^*(y)\big) = b\big(x, u_2^*\big(u_1^*(y)\big)\big),$$

d'où l'égalité $(u_1 \circ u_2)^* = u_2^* \circ u_1^*$.

De même, si $u_1,\ u_2 \in \mathscr{L}(E)$ et $\lambda \in K$, pour tous $x \in E,\ y \in F$, on a

$$\begin{aligned} b\big((u_1 + \lambda u_2)(x), y\big) &= b\big(u_1(x) + \lambda u_2(x), y\big) \\ &= b\big(u_1(x), y\big) + \lambda b\big(u_2(x), y\big) \\ &= b\big(x, u_1^*(y)\big) + \lambda b\big(x, u_2^*(y)\big) \\ &= b\big(x, u_1^*(y) + \lambda u_2^*(y)\big) \\ &= b\big(x, (u_1^* + \lambda u_2^*)(y)\big), \end{aligned}$$

d'où l'égalité $(u_1 + \lambda u_2)^* = u_1^* + \lambda u_2^*$, et la linéarité désirée. Cela achève la démonstration. □

6.4. Remarques

(1) On peut adapter la démonstration pour obtenir le résultat suivant : soit $b : E \times F \longrightarrow K$ une forme bilinéaire non nulle. Les propriétés suivantes sont équivalentes :

(i) l'application linéaire b_* est surjective ;

(ii) pour tout $u \in \mathscr{L}(E)$, il existe un endomorphisme $u^* \in \mathscr{L}(F)$ (non nécessairement unique) tel que

$$b\big(u(x), y\big) = b\big(x, u^*(y)\big), \qquad \text{pour tous } x \in E,\ y \in F.$$

L'implication (ii) $\Longrightarrow$ (i) se démontre comme précédemment. Pour démontrer (i) $\Longrightarrow$ (ii), on fixe une base $\mathbf{f} = (f_j)_{j \in J}$ de F, et pour tout indice $j \in J$, on choisit $z_j \in F$ tel que $b_*(z_j) = \theta_{f_j}$.

Il existe alors un unique $u^* \in \mathscr{L}(F)$ tel que $u^*(f_j) = z_j$, pour tout indice $j \in J$. En utilisant la bilinéarité de b, on montre que u^* vérifie les conditions souhaitées. Les détails sont laissés au lecteur.

(2) Si E et F sont de même dimension finie $n \geqslant 1$, la démonstration de l'égalité $\mathrm{Im}(u^*) = \mathrm{Ker}(u)^\perp$ est nettement plus simple.

En effet, dans ce cas, la bijectivité de b_* implique l'injectivité de $_*b$.

Mais alors, pour tout $x \in E$, on a

$$\begin{aligned} x \in {}^{\perp}\mathrm{Im}(u^*) &\iff b\big(x, u^*(y)\big) = 0, \quad \text{pour tout } y \in F \\ &\iff b\big(u(x), y\big) = 0, \quad \text{pour tout } y \in F \\ &\iff {}_*b\big(u(x)\big) = 0 \\ &\iff u(x) = 0_E \\ &\iff x \in \mathrm{Ker}(u). \end{aligned}$$

Ainsi, ${}^{\perp}\mathrm{Im}(u^*) = \mathrm{Ker}(u)$, et l'on en déduit que $\mathrm{Im}(u^*) = \mathrm{Ker}(u)^{\perp}$ par le théorème 4.5.

6.5. Définition. Soit $b : E \times F \longrightarrow K$ une forme bilinéaire.

Si $b_* : F \longrightarrow E^*$ est bijective, et si $u \in \mathscr{L}(E)$, l'endomorphisme u^* s'appelle l'endomorphisme *adjoint* à u (ou encore, ajoint de u relativement à b).

De même, si ${}_*b : E \longrightarrow F^*$ est bijective, et si $\rho \in \mathscr{L}(F)$, l'endomorphisme ${}^*\rho$ s'appelle l'endomorphisme *adjoint* à ρ.

6.6. Remarque. Si $E = F$ et b est symétrique, les deux adjoints coïncident (lorsqu'ils existent), et l'on retrouve la définition usuelle de l'adjoint d'un endomorphisme.

6.7. Corollaire. *Soit $b : E \times F \longrightarrow K$ une forme bilinéaire.*

(1) *On suppose que $b_* : F \longrightarrow E^*$ est bijective. Alors, pour tout $u \in \mathscr{L}(E)$, on a*

$$u \text{ surjective (resp. injective)} \iff u^* \text{ injective (resp. surjective)}.$$

(2) *On suppose que ${}_*b : E \longrightarrow F^*$ est bijective. Alors, pour tout $\rho \in \mathscr{L}(F)$, on a*

$$\rho \text{ surjective (resp. injective)} \iff {}^*\rho \text{ injective (resp. surjective)}.$$

Démonstration. Il suffit de démontrer le premier point.

Si u est surjective, par le théorème 6.3, on a

$$\mathrm{Ker}(u^*) = \mathrm{Im}(u)^{\perp} = E^{\perp} = \{0_F\},$$

la dernière égalité découlant de l'injectivité de b_*. Ainsi, u^* est injective. Inversement, si u^* est injective, on a $\mathrm{Im}(u)^{\perp} = \mathrm{Ker}(u^*) = \{0_F\}$ et par suite

$$\mathrm{Im}(u) = {}^{\perp}\big(\mathrm{Im}(u)^{\perp}\big) = {}^{\perp}\{0_F\} = E,$$

la première égalité découlant du théorème 4.3, b_* étant surjective. Ainsi, u est surjective.

Si u est injective, par le théorème 6.3, on a

$$\mathrm{Im}(u^*) = \mathrm{Ker}(u)^\perp = \{0_E\}^\perp = F,$$

et u^* est surjective. Inversement, si u^* est surjective, on a $\mathrm{Ker}(u)^\perp = F$, et

$$\mathrm{Ker}(u) = {}^\perp\big(\mathrm{Ker}(u)^\perp\big) = {}^\perp F = \{0_E\},$$

la dernière égalité découlant de l'injectivité de ${}_*b$, qui découle elle-même de la surjectivité de b_* (voir le théorème 4.3 ou le lemme 2.11). Ainsi, u est injective. □

On finit en donnant quelques compléments sur les endomorphismes adjoints dans le cas de la dimension finie.

6.8. Théorème. *Soit $b : E \times F \longrightarrow K$ une forme bilinéaire, où E et F sont non nuls. Les propriétés suivantes sont équivalentes :*

(1) *tout endomorphisme $u \in \mathscr{L}(E)$ et tout endomorphisme $\rho \in \mathscr{L}(F)$ admettent un unique adjoint ;*

(2) *la forme bilinéaire b est non dégénérée, et E et F sont de dimension finie.*

Dans ce cas, $\dim_K(E) = \dim_K(F)$. De plus, on a

$$^*(u^*) = u \quad \text{et} \quad (^*\rho)^* = \rho.$$

Enfin, soient $\mathbf{e}$ et $\mathbf{f}$ des bases de E et F respectivement, et soient

$$B = \mathrm{Mat}(b\,;\mathbf{e},\mathbf{f}), \quad M = \mathrm{Mat}(u\,;\mathbf{e}), \quad N = \mathrm{Mat}(\rho\,;\mathbf{f}).$$

Alors, on a

$$\mathrm{Mat}(u^*;\mathbf{f}) = B^{-1}M^tB \quad \text{et} \quad \mathrm{Mat}(^*\rho\,;\mathbf{e}) = (B^t)^{-1}N^tB^t.$$

En particulier, si $\mathbf{e}$ et $\mathbf{f}$ sont duales par rapport à b, on a

$$\mathrm{Mat}(u^*\,;\mathbf{f}) = M^t \quad \text{et} \quad \mathrm{Mat}(^*\rho\,;\mathbf{e}) = N^t.$$

Démonstration

(1) $\Longrightarrow$ (2). Les hypothèses et le théorème 6.3 impliquent que les applications b_* et ${}_*b$ sont bijectives. Par le lemme 2.11 (3), b est non dégénérée, et E et F sont de dimension finie, de même dimension.

(2) $\Longrightarrow$ (1). Si E et F sont de dimension finie et b est non dégénérée, alors ce même lemme implique que b_* et ${}_*b$ sont bijectives. On applique alors le théorème 6.3.

Dans ce cas, si $u \in \mathscr{L}(E)$, alors ${}^*(u^*) \in \mathscr{L}(E)$ est l'unique endomorphisme de E vérifiant

$$b\big({}^*(u^*)(x), y\big) = b\big(x, u^*(y)\big), \quad \text{pour tout } x \in E,\ y \in E.$$

Mais u vérifie

$$b\big(u(x), y\big) = b\big(x, u^*(y)\big), \quad \text{pour tout } x \in E,\ y \in E,$$

et par la propriété d'unicité, on en déduit que ${}^*(u^*) = u$. On procède de même pour l'autre égalité.

Gardons les notations de l'énoncé, et soit $M^* = \mathrm{Mat}(u^*; \mathbf{f})$.

D'après le lemme 5.2, l'égalité

$$b\big(u(x), y\big) = b\big(x, u^*(y)\big), \quad \text{pour tous } x \in E,\ y \in F,$$

s'écrit alors matriciellement

$$X^t M^t B Y = X^t B M^* Y, \quad \text{pour tous } X, Y \in K^n.$$

En appliquant cette égalité aux divers vecteurs de base de K^n, on obtient alors $M^t B = B M^*$. Puisque b est non dégénérée, B est inversible par la proposition 5.3, et l'on a donc $M^* = B^{-1} M^t B$. On obtient de même l'autre égalité. La dernière partie découle du fait que si $\mathbf{e}$ et $\mathbf{f}$ sont duales par rapport à b, alors $\mathrm{Mat}(b; \mathbf{e}, \mathbf{f}) = I_n$. Cela achève la démonstration. □

6.9. Exemple. Soit E un K-espace vectoriel. Puisque $(\langle\ ,\ \rangle_E)_*$ est Id_{E^*}, qui est bijective, tout $u \in \mathscr{L}(E)$ admet un unique adjoint $u^* \in \mathscr{L}(E^*)$. Cet adjoint est l'endomorphisme

$$\begin{aligned} E^* &\longrightarrow E^* \\ \varphi &\longmapsto \varphi \circ u. \end{aligned}$$

En effet, $u^* \in \mathscr{L}(E^*)$ est l'unique endomorphisme de E^* satisfaisant

$$\varphi\big(u(x)\big) = \big(u^*(\varphi)\big)(x), \quad \text{pour tous } x \in E,\ \varphi \in E^*.$$

Cet endomorphisme s'appelle la *transposée* de u ou l'endomorphisme transposé [10] de u .

Supposons que E soit de dimension finie. Soit $\mathbf{e}$ une base de E et soit $\mathbf{e}^*$ la base duale. Posons $M = \mathrm{Mat}(u; \mathbf{e})$. D'après ce qui précède, on a

$$\mathrm{Mat}(u^*; \mathbf{e}^*) = M^t.$$

De plus, on a

$$\mathrm{Mat}\big({}^*(u^*); \mathbf{e}\big) = (M^t)^t,$$

et comme ${}^*(u^*) = u$, on obtient $(M^t)^t = M$.

10. Une légère généralisation de cette notion est proposée en exercice.

La dualité fournit ainsi une explication conceptuelle à l'égalité $(M^t)^t = M$ lorsque M est une matrice carrée.

De plus, l'égalité $\mathrm{rg}(u^*) = \mathrm{rg}(u)$ pour tout $u \in \mathscr{L}(E)$ donnée par le théorème 6.3 se récrit

$$\mathrm{rg}(M^t) = \mathrm{rg}(M), \quad \text{pour tout } M \in \mathrm{M}_n(K).$$

Enfin, définissons une application bilinéaire

$$\begin{aligned} b_0 \colon K^n \times K^n &\longrightarrow K \\ (X, Y) &\longmapsto X^t Y. \end{aligned}$$

L'égalité $\mathrm{Im}(u^*) = \mathrm{Ker}(u)^\perp$ donne alors

$$\mathrm{Im}(M^t) = \mathrm{Ker}(M)^\perp, \quad \text{pour tout } M \in \mathrm{M}_n(K),$$

où l'orthogonal est défini par rapport à b_0. Autrement dit,

$$\mathrm{Im}(M^t) = \{X \in K^n \mid X^t Y = 0, \quad \text{pour tout } Y \in \mathrm{Ker}(M)\}.$$

De même, l'égalité $\mathrm{Ker}(u^*) = \mathrm{Im}(u)^\perp$ donne

$$\mathrm{Ker}(M^t) = \{X \in K^n \mid X^t Y = 0, \quad \text{pour tout } Y \in \mathrm{Im}(M)\}.$$

7. Exercices

E1. Soit $\varphi \in E^*$ une forme linéaire non nulle. Montrer que l'application

$$\begin{aligned} b \colon E \times E &\longrightarrow K \\ (x, y) &\longmapsto \varphi(x)\varphi(y) \end{aligned}$$

est une forme bilinéaire. Est-elle non dégénérée ?

E2. Soit K un corps de caractéristique différente de 2, soit E un K-espace vectoriel de dimension finie et soit $b : E \times E \longrightarrow K$ une forme bilinéaire antisymétrique, i.e. telle que

$$b(y, x) = -b(x, y), \quad \text{pour tous } x, y \in E.$$

Montrer que si b est non dégénérée, alors $\dim_K(E)$ est paire.

E3. Soit K un corps.

a. Donner l'exemple d'une forme bilinéaire $b : K^2 \times K^2 \longrightarrow K$ et de deux vecteurs $x, y \in K^2$ tels que $x \perp_b y$ mais $y \not\perp_b x$.

b. On suppose que $b : E \times E \longrightarrow K$ est symétrique ou antisymétrique. Vérifier que pour tous $x, y \in E$, on a $x \perp_b y$ si, et seulement si, $y \perp_b x$.

On s'intéresse maintenant à la réciproque.

c. Soient E et F deux K-espaces vectoriels, et soient $u, v \in \mathscr{L}(E,F)$ deux applications linéaires. On suppose que pour tout vecteur $x \in E$, il existe $\lambda_x \in K^\times$ tel que $u(x) = \lambda_x v(x)$.

(i) Vérifier que u et v ont même noyau et même image.

(ii) On considère les isomorphismes

$$\overline{u} : E/\mathrm{Ker}(u) \xrightarrow{\sim} \mathrm{Im}(u) \quad \text{et} \quad \overline{v} : E/\mathrm{Ker}(v) \xrightarrow{\sim} \mathrm{Im}(v)$$

induits par u et v. Montrer que $\varphi = \overline{v}^{-1} \circ \overline{u}$ est un endomorphisme de $E/\mathrm{Ker}(u)$ vérifiant

$$\varphi(\overline{x}) = \lambda_x \cdot \overline{x}, \quad \text{pour tout } \overline{x} \in E/\mathrm{Ker}(u).$$

En déduire que φ est une homothétie.

(iii) Montrer alors qu'il existe $\lambda \in K$, tel que $u(x) = \lambda v(x)$, pour tout vecteur $x \in E$.

d. On suppose que $b : E \times E \longrightarrow K$ est une forme bilinéaire telle pour tous $x, y \in E$, on a

$$x \perp_b y \iff y \perp_b x.$$

(i) Montrer que pour tout $x \in E$, les formes linéaires $b_*(x)$ et ${}_*b(x)$ ont même noyau.

(ii) En déduire **soigneusement** que pour tout $x \in E$, il existe $\lambda_x \in K^\times$, tel que ${}_*b(x) = \lambda_x b_*(x)$.

(iii) Montrer qu'il existe $\lambda \in K$, tel que $b(x,y) = \lambda b(y,x)$, pour tous vecteurs $x, y \in E$.

(iv) Montrer que b est symétrique ou antisymétrique (on prendra soin de traiter le cas où b est nulle à part).

E4. Soient $a_0, \ldots, a_n \in \mathbb{R}$ des réels distincts.

Montrer qu'il existe des réels $b_0, \ldots, b_n \in \mathbb{R}$ uniques tels que, pour tout polynôme $P \in \mathbb{R}_n[X]$, on a

$$\int_0^1 P(t)\mathrm{d}t = b_0 P(a_0) + \cdots + b_n P(a_n).$$

E5. Soit $E = \mathbb{R}_n[X]$, et soient $a_0, \ldots, a_n \in \mathbb{R}$ des réels **distincts**. Pour tout $j \in [\![0, n]\!]$, soit

$$\begin{aligned} f_j : E &\longrightarrow \mathbb{R} \\ P &\longmapsto P(a_j). \end{aligned}$$

Montrer que $\mathbf{f} = (f_0, \ldots, f_n)$ est une base de E^*.

Montrer également que la base duale de $\mathbf{f}$ pour le crochet de dualité est $\mathbf{e} = (L_0, \ldots, L_n)$, où L_i est défini par

$$L_i = \prod_{j \neq i} \frac{X - a_j}{a_i - a_j}.$$

E6. Soit $E = \mathbb{R}_n[X]$, et soit $\mathbf{e} = (1, X, \ldots, X^n)$. Pour $j \in [\![0, n]\!]$, posons

$$\begin{aligned} f_j : E &\longrightarrow \mathbb{R} \\ P &\longmapsto \frac{P^{(j)}(0)}{j!}. \end{aligned}$$

Montrer que $\mathbf{f} = (f_0, \ldots, f_n)$ est la base duale de $\mathbf{e}$ par rapport au crochet de dualité.

E7. Soit $P \in \mathbb{R}[X]$ un polynôme de degré d fixé. On se propose dans cet exercice d'exprimer la somme

$$S_n(P) = \sum_{k=0}^{n} P(k)$$

comme un polynôme de degré $\leqslant d+1$ en $n+1$. On note $E = \mathbb{R}_{d+1}[X]$.

a. Pour tout $i \in [\![0, d+1]\!]$, on pose

$$P_i = \frac{X(X-1)\cdots(X-i+1)}{i!}.$$

Vérifier que la famille $(P_0, \ldots, P_{d+1})$ est une base de E.

b. Pour tout $P \in E$, on pose $\Delta(P) = P(X+1) - P(X)$.

Vérifier que $\Delta \in \mathscr{L}(E)$, et calculer $\Delta(P_i)$, pour tout $i \in [\![0, d+1]\!]$.

c. On pose $\Delta^0 = \mathrm{Id}_E$ et $\Delta^i = \Delta \circ \cdots \circ \Delta$, $i \geqslant 1$. Vérifier que la famille

$$\left(\Delta^0(\_)(0), \ldots, \Delta^{d+1}(\_)(0)\right)$$

est la base duale de $(P_0, \ldots, P_{d+1})$ (pour le crochet de dualité)

d. Soit $Q = \sum_{i=0}^{d} \Delta^i(P)(0) P_{i+1}$.

Déduire des questions précédentes que $P = \Delta(Q)$, puis l'égalité

$$S_n(P) = Q(n+1).$$

e. Calculer $S_n(X)$, $S_n(X^2)$ et $S_n(X^3)$.

f. Montrer que pour tout $P \in \mathbb{R}[X]$, et tout $k \geqslant 0$, on a

$$\Delta^k(P) = \sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} P(X+i).$$

g. En déduire que pour tout $d \geqslant 0$, on a

$$S_d = \sum_{k=0}^{d} \Big(\sum_{i=0}^{k} (-1)^{k-i} \binom{k}{i} i^d \Big) \binom{n+1}{k+1}.$$

E8. Soit K un corps et soit $n \geqslant 1$ un entier.

a. Montrer que l'application

$$\begin{array}{rccc} \mathcal{T} : & \mathrm{M}_n(K) \times \mathrm{M}_n(K) & \longrightarrow & K \\ & (A, B) & \longmapsto & \mathrm{tr}(AB) \end{array}$$

est une forme bilinéaire symétrique non dégénérée.

b. Soit H un hyperplan de $\mathrm{M}_n(K)$. Justifier qu'il existe $A \in \mathrm{M}_n(K)$ telle H soit le noyau de la forme linéaire

$$\begin{array}{rccc} \varphi : & \mathrm{M}_n(K) & \longrightarrow & K \\ & M & \longmapsto & \mathrm{tr}(AM). \end{array}$$

c. En déduire que tout hyperplan de $\mathrm{M}_n(K)$ contient une matrice inversible.

Indication. On pourra utiliser la question précédente et se ramener au cas où $A = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$.

E9. Soit E un K-espace vectoriel de dimension finie, et soit $u \in \mathscr{L}(E)$. Montrer qu'il existe une forme bilinéaire symétrique non dégénérée pour laquelle u est autoadjoint (i.e. $u^* = u$).

Indication. On pourra utiliser l'exercice XXXI-6.

E10. On se donne un tableau électrique où sont disposées n ampoules, chaque ampoule pouvant être, soit allumée, soit éteinte. L'installation électrique est soumise aux règles suivantes :

(i) Actionner un interrupteur change l'état de l'ampoule correspondante et de certaines autres ampoules, et qui sont toujours les mêmes à chaque nouvelle action sur ce même interrupteur. En particulier, actionner deux fois de suite le même interrupteur ne change rien.

(ii) Pour tous $i, j \in [\![1, n]\!]$, si actionner le j-ème interrupteur change (entre autres) l'état de la i-ème ampoule, alors actionner le i-ème interrupteur change (entre autres) l'état de la j-ème ampoule.

Une *configuration d'ampoules* est un état du tableau obtenu en actionnant successivement certains interrupteurs à partir d'un tableau complètement éteint.

On numérote les ampoules de 1 à n. À toute configuration d'ampoules, on associe un vecteur $X \in \mathbb{F}_2^n$, appelé *vecteur d'état*, dont la i-ème coordonnée vaut $\overline{1}$ si, et seulement si, la i-ème ampoule est allumée.

On appelle *i-ème configuration élémentaire* la configuration obtenue en actionnant l'interrupteur i à partir d'un tableau complètement éteint. On note X_i le vecteur d'état correspondant.

La *matrice du jeu d'ampoules* est par définition la matrice $A \in \mathrm{M}_n(\mathbb{F}_2)$ dont les colonnes sont $X_1, \ldots, X_n$.

a. Soit $X \in \mathbb{F}_2^n$. Montrer que X est le vecteur d'état d'une configuration d'ampoules si, et seulement si, $X \in \mathrm{Im}(A)$.

b. Montrer que A est symétrique, et possède des $\overline{1}$ sur sa diagonale.

c. Montrer que le vecteur $U \in \mathbb{F}_2^n$ dont toutes les coordonnées sont des $\overline{1}$ appartient à $\mathrm{Im}(A)$.

Indication. Utiliser l'exemple 6.9.

d. En déduire que, pour n'importe quel jeu d'ampoules, on peut toujours éteindre un tableau dont toutes les ampoules sont allumées.

Le lecteur intéressé trouvera d'autres résultats sur les jeux d'ampoules dans [10].

E11. Soit K un corps possédant au moins trois éléments, et soit E un K-espace vectoriel de dimension $n \geqslant 1$. On se propose de démontrer que tout endomorphisme de E est somme de deux automorphismes de E.

a. Soient $e_1, \ldots, e_r \in E$ des vecteurs de E linéairement indépendants, et soient $\varphi_1, \ldots, \varphi_r \in E^*$. Montrer que l'endomorphisme

$$\begin{aligned} E &\longrightarrow E \\ x &\longmapsto \sum_{i=1}^{r} \varphi_i(x) e_i \end{aligned}$$

est de rang r si, et seulement si, $\varphi_1, \ldots, \varphi_r$ sont linéairement indépendantes.

Indication. Utiliser le théorème du rang.

b. Soit $u \in \mathscr{L}(E)$ un endomorphisme de rang r.

(1) Montrer qu'il existe $e_1, \dots, e_r \in E$ linéairement indépendants, et des formes linéaires $\varphi_1, \dots, \varphi_r \in E^*$ linéairement indépendantes telles que

$$u(x) = \sum_{i=1}^{r} \varphi_i(x) e_i, \quad \text{pour tout } x \in E.$$

(2) On complète la famille $(e_1, \dots, e_r)$ en une base $(e_1, \dots, e_n)$ de E, ainsi que la famille $(\varphi_1, \dots, \varphi_r)$ en une base $(\varphi_1, \dots, \varphi_n)$ de E^*.

Soit $\lambda \in K \setminus \{0, 1\}$, et soient u_1, $u_2 \in \mathscr{L}(E)$ définis par

$$\begin{aligned} u_1 : E &\longrightarrow E \\ x &\longmapsto \lambda \sum_{i=1}^{r} \varphi_i(x) e_i + \sum_{i=r+1}^{n} \varphi_i(x) e_i \end{aligned}$$

et

$$\begin{aligned} u_1 : E &\longrightarrow E \\ x &\longmapsto (1-\lambda) \sum_{i=1}^{r} \varphi_i(x) e_i - \sum_{i=r+1}^{n} \varphi_i(x) e_i. \end{aligned}$$

Montrer que u_1 et u_2 sont des automorphismes de E, et conclure.

E12. Soit $b : E \times F \longrightarrow K$ une forme bilinéaire non dégénérée, où E et F sont deux K-espaces vectoriels de dimension n. Ainsi, b_* et ${}_*b$ sont bijectives.

On se propose ici de démontrer de manière directe la correspondance entre les sous-espaces de E et ceux de F, sans utiliser les résultats démontrés dans ce chapitre.

a. Soit V un sous-espace de E. Montrer que toute forme linéaire sur V se prolonge en une forme linéaire sur E. En déduire que l'application

$$\begin{aligned} F &\longrightarrow V^* \\ y &\longmapsto b_*(y)_{|V} \end{aligned}$$

est surjective, de noyau $V^\perp$.

b. En déduire que $\dim_K(V^\perp) = n - \dim_K(V)$.

c. Montrer que $\dim_K \left({}^\perp(V^\perp)\right) = \dim_K(V)$, puis que ${}^\perp(V^\perp) = V$, et conclure.

E13. Soit $b : E \times F \longrightarrow K$ une forme bilinéaire, et soient V_1, V_2 deux sous-espaces de E.

a. Montrer que $(V_1 + V_2)^\perp = V_1^\perp \cap V_2^\perp$.

b. Montrer que $V_1^\perp + V_2^\perp \subset (V_1 \cap V_2)^\perp$.

c. On suppose que $b_* : F \longrightarrow E^*$ est surjective. On se propose de montrer l'égalité

$$V_1^\perp + V_2^\perp = (V_1 \cap V_2)^\perp.$$

(i) Pour $i = 1, 2$, soit V_i' un supplémentaire de $V_1 \cap V_2$ dans V_i. Montrer que $V_1 \cap V_2$, V_1' et V_2' sont en somme directe. Soit V_3' un supplémentaire de $V_1 \cap V_2 \oplus V_1' \oplus V_2'$ dans E.

Soit $y \in (V_1 \cap V_2)^\perp$.

(ii) Montrer qu'il existe deux formes linéaires φ_1, $\varphi_2 \in E^*$ telles que

$$\varphi_1(x) = \begin{cases} 0 & \text{si} \quad x \in V_1 \\ b(x,y) & \text{si} \quad x \in V_2' \oplus V_3' \end{cases}$$

et

$$\varphi_2(x) = \begin{cases} 0 & \text{si} \quad x \in V_2 \oplus V_3' \\ b(x,y) & \text{si} \quad x \in V_1'. \end{cases}$$

(iii) Montrer que $\varphi_1 + \varphi_2 = b_*(y)$.

(iv) Soient y_1, $y_2 \in F$ tels que $\varphi_1 = b_*(y_1)$ et $\varphi_2 = b_*(y_2)$.

Montrer que $y_1 \in V_1^\perp$, $y_2 \in V_2^\perp$ et que $y - y_1 - y_2 \in \mathrm{Ker}(b_*)$.

(v) En déduire que $y \in V_1^\perp + V_2^\perp$ et conclure.

d. On suppose que b_* est bijective. Montrer que $E = V_1 \oplus V_2$ si, et seulement si, $F = V_1^\perp \oplus V_2^\perp$.

E14. Soit $u : E \longrightarrow F$ une application linéaire. L'application *transposée* est l'application linéaire

$$\begin{aligned} u^t : F^* &\longrightarrow E^* \\ \varphi &\longmapsto \varphi \circ u. \end{aligned}$$

a. Montrer les égalités

$$\mathrm{Im}(u^t) = \mathrm{Ker}(u)^\circ \quad \text{et} \quad \mathrm{Ker}(u^t) = \mathrm{Im}(u)^\circ.$$

b. Montrer que u^t est de rang fini si, et seulement si, u est de rang fini. Montrer que dans ce cas, u^t et u ont même rang.

c. Montrer que u^t est injective (resp. surjective, bijective) si, et seulement si, u est surjective (resp. injective, bijective).

Lorsque u est bijective, calculer $(u^t)^{-1}$ en fonction de u^{-1}.

d. On suppose dans cette question que E et F sont de dimension finie. Soient $\mathbf{e}$ et $\mathbf{f}$ des bases de E et F respectivement.

Comparer $\mathrm{Mat}(u\,; \mathbf{e}, \mathbf{f})$ et $\mathrm{Mat}(u^t\,; \mathbf{f}^*, \mathbf{e}^*)$.

E15. Soit E un K-espace vectoriel de dimension quelconque. On se propose dans cet exercice de démontrer directement l'égalité $({}^\circ W)^\circ = W$ pour tout sous-espace W de E^* de dimension finie.

a. Montrer qu'établir le résultat voulu revient à démontrer l'énoncé suivant : pour toute famille $(\varphi_1, \dots, \varphi_k)$ de formes linéaires sur E linéairement indépendantes, et pour tout $\varphi \in E^*$, on a

$$\bigcap_{i=1}^{k} \mathrm{Ker}(\varphi_i) \subset \mathrm{Ker}(\varphi) \iff \varphi \in \mathrm{Vect}_K(\varphi_1, \dots, \varphi_k).$$

On se propose de montrer le résultat ci-dessus par récurrence sur k.

b. Régler le cas $k = 0$.

c. On suppose avoir démontré le résultat pour toute famille libre de k formes linéaires. Soient $\varphi_1, \dots, \varphi_{k+1} \in E^*$ $k+1$ formes linéaires linéairement indépendantes sur K.

(1) Pour tout $i \in [\![1, k+1]\!]$, montrer par l'absurde l'existence d'un vecteur $x_i' \in \bigcap_{j \neq i} \mathrm{Ker}(\varphi_j)$ tel que $\varphi_i(x_i') \neq 0$.

(2) En déduire l'existence de vecteurs $x_1, \dots, x_{k+1} \in E$ tels que

$$\varphi_j(x_i) = \delta_{ij} \text{ pour tous } i, j \in [\![1, k+1]\!].$$

d. Avec les notations de la question précédente, montrer que l'on a

$$E = \mathrm{Vect}_K(x_1, \dots, x_{k+1}) \oplus \bigcap_{i=1}^{k+1} \mathrm{Ker}(\varphi_i).$$

e. Conclure alors que le résultat voulu est vrai pour $k+1$ formes linéaires linéairement indépendantes.

f. Déduire également de la démonstration précédente le résultat suivant. Soient $\varphi_1, \dots, \varphi_r \in E^*$ des formes linéaires linéairement indépendantes ; alors, il existe des vecteurs $x_1, \dots, x_r \in E$ tels que

$$\varphi_i(x_j) = \delta_{ij} \text{ pour tous } i, j \in [\![1, r]\!].$$

De plus, si $\psi \in E^*$ vérifie $\bigcap_{i=1}^{r} \mathrm{Ker}(\varphi_i) \subset \mathrm{Ker}(\psi)$, alors $\psi = \sum_{i=1}^{r} \psi(x_i)\varphi_i$.

g. Lorsque E est de dimension finie, proposer une méthode pour calculer explicitement de tels vecteurs $x_1, \dots, x_r \in E$.

Chapitre XXXIII

Valeurs propres de matrices symétriques

1. Introduction

Le but de ce chapitre est d'utiliser les outils de théorie des corps et d'algèbre linéaire introduits dans les chapitres précédents pour étudier de plus près les propriétés des matrices symétriques sur un corps quelconque.

Le point de départ est le suivant. Soit $M \in \mathrm{M}_n(\mathbb{R})$ une matrice symétrique réelle. Le théorème spectral nous dit que toutes les valeurs propres de M sont réelles et que M est diagonalisable. Inversement, un réel arbitraire est valeur propre d'une matrice symétrique réelle, puisqu'il suffit de prendre une matrice diagonale appropriée.

Ainsi, l'ensemble des réels qui sont valeurs propres d'une matrice symétrique réelle est exactement $\mathbb{R}$, ce qui n'est pas très passionnant en soi. Néanmoins, pour corser un peu l'affaire, on peut se demander ce qui se passe si l'on remplace $\mathbb{R}$ par autre chose. Plus précisément, on peut se poser la question suivante. Pour k un corps quelconque, peut-on décrire l'ensemble des valeurs propres d'une matrice symétrique à coefficients dans k ?

Clairement, les valeurs propres de matrices symétriques à coefficients dans k sont algébriques sur k, puisqu'une valeur propre est racine du polynôme caractéristique.

La véritable question est donc la suivante. Pour k un corps, un élément quelconque de $\overline{k}$ est-il valeur propre d'une matrice symétrique à coefficients dans k ?

Pour nous simplifier un peu la vie, on peut déjà se demander ce qui se passe lorsque $k = \mathbb{Q}$. Pour donner un élément de réponse à cette question, notons qu'une matrice symétrique à coefficients rationnels peut se voir comme une matrice symétrique à coefficients réels. En particulier, elle est diagonalisable sur $\mathbb{R}$, et son polynôme minimal est donc scindé sur $\mathbb{R}$. Cela conduit à la définition suivante.

1.1. Définition. On dit qu'un nombre complexe $\alpha \in \mathbb{C}$ est *totalement réel* s'il existe un polynôme unitaire $f \in \mathbb{Q}[X]$ scindé sur $\mathbb{R}$ tel que $f(\alpha) = 0$. En particulier, un nombre totalement réel est un réel algébrique sur $\mathbb{Q}$.

De manière équivalente[1], un nombre complexe $\alpha \in \mathbb{C}$ est totalement réel s'il est algébrique sur $\mathbb{Q}$, et son polynôme minimal $\mu_{\alpha,\mathbb{Q}}$ est scindé sur $\mathbb{R}$.

1.2. Exemple. Si $d \in \mathbb{Q}$ est un rationnel strictement positif, $\sqrt{d}$ est totalement réel, puisqu'il est racine de $X^2 - d = (X - \sqrt{d})(X + \sqrt{d})$.

Les valeurs propres de matrices symétriques à coefficients dans $\mathbb{Q}$ sont donc à chercher parmi les nombres totalement réels.

La première question à se poser est de savoir si tout réel est totalement réel. Le résultat suivant montre que la réponse est négative.

1.3. Proposition. *Soit $d \in \mathbb{Q}$ un rationnel positif. On suppose que d n'est pas le cube d'un rationnel. Alors, le réel $\alpha = \sqrt[3]{d}$ n'est pas totalement réel. En particulier, α n'est pas valeur propre d'une matrice symétrique à coefficients rationnels.*

Démonstration. Par hypothèse sur d, $X^3 - d \in \mathbb{Q}[X]$ n'a pas racines rationnelles. Il est donc irréductible sur $\mathbb{Q}$, étant donné qu'il est de degré 3. Comme il est unitaire et annule α, on a $\mu_{\alpha,\mathbb{Q}} = X^3 - d$. Mais, les racines de ce polynôme sont α, $j\alpha$ et $j^2\alpha$. Ainsi, $\mu_{\alpha,\mathbb{Q}}$ n'est pas scindé sur $\mathbb{R}$. Par conséquent, α n'est pas totalement réel, et ne peut donc être valeur propre d'une matrice symétrique à coefficients rationnels. □

Notation. Dans la suite, si k est un corps, $\mathscr{S}_n(k)$ désignera l'ensemble des matrices symétriques de $\mathrm{M}_n(k)$.

Revenons à notre problème initial, à savoir : tout nombre totalement réel est-il valeur propre d'une matrice symétrique à coefficients rationnels ?

Pour bien faire comprendre la difficulté de la question, commençons par étudier brièvement le cas de $\sqrt{d}$. Si d est le carré d'un rationnel, $\sqrt{d}$ est rationnel, et $\sqrt{d}$ est bien entendu valeur propre de la matrice $(\sqrt{d}) \in \mathscr{S}_1(\mathbb{Q})$.

1. La démonstration de cette équivalence est laissée à la sagacité du lecteur.

On suppose dorénavant que d n'est pas le carré d'un rationnel. Pour essayer de déterminer si $\sqrt{d}$ est valeur propre d'une matrice symétrique à coefficients rationnels, on peut commencer par examiner le cas des matrices de taille 2.

Supposons que $\sqrt{d}$ soit valeur propre de $M = \begin{pmatrix} u & v \\ v & w \end{pmatrix} \in \mathscr{S}_2(\mathbb{Q})$. Alors, la matrice réelle $M - \sqrt{d}\mathrm{I}_2$ n'est pas inversible, donc est de déterminant nul. On a alors

$$(u - \sqrt{d})(w - \sqrt{d}) - v^2 \;=\; uw - v^2 + d - (u+w)\sqrt{d} \;=\; 0.$$

Si $u+w \neq 0$, on obtient $\sqrt{d} \in \mathbb{Q}$, puis que d est le carré d'un rationnel, d'où une contradiction. Ainsi, $u + w = 0$, et donc $-u^2 - v^2 + d = 0$. Autrement dit, d est somme de deux carrés dans $\mathbb{Q}$.

Réciproquement, si $d = u^2 + v^2$, avec u, $v \in \mathbb{Q}$, alors il est facile de voir que le polynôme caractéristique de la matrice symétrique $\begin{pmatrix} u & -v \\ -v & -u \end{pmatrix}$ est $X^2 - d$, ce qui démontre que $\sqrt{d}$ est valeur propre d'une matrice de $\mathrm{M}_2(\mathbb{Q})$. On a donc établi le résultat suivant.

1.4. Proposition. *Soit $d \in \mathbb{Q}$ un rationnel positif. On suppose que d n'est pas le carré d'un rationnel.*

Alors, $\sqrt{d}$ est valeur propre d'une matrice de $\mathscr{S}_2(\mathbb{Q})$ si, et seulement si, d est une somme de deux carrés dans $\mathbb{Q}$.

1.5. Exemple. Le résultat précédent montre que $\sqrt{3}$ n'est pas valeur propre d'une matrice de $\mathscr{S}_2(\mathbb{Q})$.

Il suffit en effet de se convaincre que 3 n'est pas somme de deux carrés dans $\mathbb{Q}$.

Supposons au contraire que $3 = u^2 + v^2$, u, $v \in \mathbb{Q}$. En réduisant au même dénominateur, et en multipliant par le carré d'un dénominateur commun, on obtient que $3c^2 = a^2 + b^2$, $a, b, c \in \mathbb{Z}$, $c \neq 0$. Si a et b sont tous deux divisibles par 3, alors $3^2 \mid 3c^2$. Mais alors, $3 \mid c^2$, puis $3 \mid c$ et enfin $3^2 \mid c$. Quitte à diviser a, b, c par 3 suffisamment de fois, on peut donc supposer que $3 \nmid a$ par exemple. En réduisant modulo 3, on obtient $a^2 + b^2 \equiv 0\ [3]$.

Puisque a n'est pas égal à 0 modulo 3, on a $a \equiv \pm 1\ [3]$, et ainsi $b^2 \equiv -1\ [3]$. Or, cette égalité ne peut avoir lieu, comme on le constate en remplaçant b par 0, 1 et 2.

L'exemple ci-dessus montre seulement que $\sqrt{3}$ n'est pas valeur propre d'une matrice de $\mathscr{S}_2(\mathbb{Q})$. Il n'exclut pas la possibilité que $\sqrt{3}$ soit valeur propre

d'une matrice de $\mathscr{S}_n(\mathbb{Q})$ avec $n > 2$. D'ailleurs, on peut vérifier que le polynôme caractéristique de la matrice $\begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix}$ est égal à $X(X^2-3)$, ce qui démontre que $\sqrt{3}$ est bien valeur propre d'une matrice de $\mathscr{S}_3(\mathbb{Q})$.

Bref, ce n'est pas si simple...

Pour mettre tout de suite fin à ce suspense insoutenable, donnons tout de suite la réponse au problème initial : tout nombre totalement réel est effectivement valeur propre d'une matrice symétrique à coefficients rationnels.

1.6. Théorème. *Tout nombre totalement réel de degré $n \geqslant 1$ sur $\mathbb{Q}$ est valeur propre d'une matrice de $\mathscr{S}_{4n}(\mathbb{Q})$.*

1.7. Remarque. Le cas de $\sqrt{3}$ montre que ce résultat n'est pas optimal. Nous reviendrons sur ce point ultérieurement.

Pour conclure cette longue introduction, remarquons que l'ensemble des valeurs propres des matrices symétriques dépend fortement du corps k choisi. Par exemple, si $d \in \mathbb{Q}$ n'est pas le cube d'un rationnel, la proposition 1.3 montre que $\sqrt[3]{d}$ n'est pas valeur propre d'une matrice symétrique à coefficients rationnels. En revanche, on constate que le polynôme caractéristique de la matrice

$$\begin{pmatrix} 0 & 1 & i \\ 1 & \dfrac{d}{4} & -\dfrac{id}{4} \\ i & -\dfrac{id}{4} & -\dfrac{d}{4} \end{pmatrix}$$

est $X^3 - d$, ce qui démontre que $\sqrt[3]{d}$ est valeur propre d'une matrice symétrique à coefficients dans $\mathbb{Q}(i)$.

Nous démontrerons en fait le résultat suivant.

1.8. Théorème. *Soit k un corps de caractéristique différente de 2. On suppose que $-1 \in k^{\times 2}$.*

Alors, tout élément algébrique de degré n sur k est valeur propre d'une matrice de $\mathscr{S}_n(k)$.

1.9. Remarque. Ce résultat est optimal. En effet, si $\alpha \in \overline{k}$ est valeur propre d'une matrice symétrique $M \in \mathrm{M}_m(k)$, alors le polynôme caractéristique de M est un polynôme de $k[X]$ de degré $m \geqslant 1$ annulant α, et donc $m \geqslant n$ par définition du degré de α. Autrement dit, n est la taille minimale d'une matrice symétrique dont α est valeur propre.

Le but de ce qui suit est de donner une explication conceptuelle aux résultats précédents.

Nous allons en fait nous intéresser à un problème un tout petit peu différent, mais néanmoins lié : étant donné un polynôme unitaire $f \in k[X]$, à quelles conditions f est-il le polynôme minimal d'une matrice symétrique à coefficients dans k ?

Résoudre cette question nous permettra de répondre à la question initiale. En effet, rappelons que le polynôme minimal et le polynôme caractéristique d'une matrice ont mêmes racines. En particulier, l'ensemble des valeurs propres des matrices symétriques à coefficients dans k sont exactement l'ensemble des racines des polynômes minimaux de telles matrices.

2. La méthode générale

Nous allons maintenant expliquer la méthode générale que nous utiliserons dans la suite.

Soit $f \in k[X]$ un polynôme unitaire de degré $n \geqslant 1$. On rappelle que la loi externe

$$\begin{aligned} k \times k[X]/(f) &\longrightarrow k[X]/(f) \\ (\lambda, \overline{P}) &\longmapsto \lambda \cdot \overline{P} = \overline{\lambda P} \end{aligned}$$

est bien définie, et induit sur l'anneau commutatif $k[X]/(f)$ une structure de k-algèbre associative, commutative et unitaire.

De plus, la famille $(\overline{1}, \overline{X}, \ldots, \overline{X}^{n-1})$ est une k-base de $k[X]/(f)$. En particulier, $k[X]/(f)$ est de dimension n (cf. proposition XXI-1.9).

L'idée qui va tout faire fonctionner est due à Bender[2] (cf. [13]), et part d'un constat très simple.

Pour tout $x \in k[X]/(f)$, l'application

$$\begin{array}{rccc} \ell_x : & k[X]/(f) & \longrightarrow & k[X]/(f) \\ & y & \longmapsto & xy \end{array}$$

est k-linéaire, et pour toute forme linéaire $s : k[X]/(f) \longrightarrow k$, on a

$$s\big(\ell_x(y_1)y_2\big) \;=\; s(xy_1y_2) \;=\; s\big(y_1(xy_2)\big) \;=\; s\big(y_1\ell_x(y_2)\big),$$

pour tous $y_1, y_2 \in k[X]/(f)$. Or, on a le résultat suivant, qui découle immédiatement du théorème XXXII-6.3, en remarquant qu'une base orthonormée pour une forme bilinéaire symétrique est autoduale.

2. Le mathématicien américain, pas le robot de Futurama !

2.1. Lemme. *Soit $b : V \times V \longrightarrow k$ une forme bilinéaire symétrique, où V est un k-espace vectoriel de dimension finie. Supposons que V possède une base $\mathscr{B} = (e_1, \ldots, e_n)$ b-orthonormée.*

Soit $u \in \mathscr{L}(V)$ un endomorphisme vérifiant

$$b\big(u(x), y\big) = b\big(x, u(y)\big), \qquad \textit{pour tous } x, y \in V.$$

Alors, $\mathrm{Mat}(u\,;\mathscr{B}) \in \mathrm{M}_n(k)$ *est symétrique.*

L'idée est donc de trouver une forme linéaire $s : E \longrightarrow k$ telles que E possède une base orthonormée $\mathscr{B}$ pour la forme bilinéaire symétrique

$$\begin{aligned} b_s : E \times E &\longrightarrow k \\ (x, y) &\longmapsto s(xy). \end{aligned}$$

Si de plus E contient un élément x tel que le polynôme minimal de ℓ_x soit égal à f, alors $\mathrm{Mat}(\ell_x\,;\mathscr{B})$ sera une matrice symétrique de polynôme minimal f.

Or, on a le lemme suivant.

2.2. Lemme. *Soit k un corps arbitraire, et soit $f \in k[X]$ un polynôme unitaire de degré $n \geqslant 1$. Posons $x = \overline{X} \in k[X]/(f)$.*

Alors, le polynôme minimal de $\ell_x \in \mathscr{L}\big(k[X]/(f)\big)$ est f.

Démonstration. De simples calculs montrent les deux propriétés suivantes :

(i) pour tout $y \in k[X]/(f)$, on a $\ell_y = 0$ si, et seulement si, $y = \overline{0}$;

(ii) pour tout $P \in k[X]$, on a $\ell_{\overline{P}} = P(\ell_x)$.

Pour tout $P \in k[X]$, on a alors

$$P(\ell_x) = 0 \iff \ell_{\overline{P}} = 0 \iff \overline{P} = \overline{0} \iff f \mid P,$$

ce qui démontre le résultat voulu. □

Résumons : si l'on dispose d'une k-forme linéaire $s : k[X]/(f) \longrightarrow k$ telle que $k[X]/(f)$ possède une b_s-orthonormée $\mathscr{B}$, alors la matrice $\mathrm{Mat}(\ell_{\overline{X}}\,;\mathscr{B})$ sera symétrique, de polynôme minimal f.

Hélas, cela n'est pas possible en général. En effet, si l'on pouvait appliquer les considérations précédentes avec $f = X^2 - 3 \in \mathbb{Q}[X]$, on obtiendrait une matrice de $\mathscr{S}_2(\mathbb{Q})$ de polynôme minimal $X^2 - 3$. En particulier, $\sqrt{3}$ serait valeur propre d'une matrice de $\mathscr{S}_2(\mathbb{Q})$. Mais c'est impossible d'après l'exemple 1.5.

Néanmoins, on a le résultat suivant.

2.3. Proposition. *Soit k un corps arbitraire, et soient $f_1, \dots, f_r \in k[X]$ des polynômes unitaires de degrés respectifs $n_1, \dots, n_r \geqslant 1$.*

Enfin, soit $E = k[X]/(f_1) \times \cdots \times k[X]/(f_r)$ et soit $x_j = \overline{X} \in k[X]/(f_j)$. On suppose qu'il existe une forme linéaire $s : E \longrightarrow k$ telle que E possède une base $\mathscr{B}$ qui est orthonormée pour la forme bilinéaire symétrique

$$\begin{aligned} b_s : E \times E &\longrightarrow k \\ (x, y) &\longmapsto s(xy). \end{aligned}$$

Si $x = (x_1, \dots, x_r)$, la matrice $M = \mathrm{Mat}(\ell_x\,; \mathscr{B}) \in \mathrm{M}_n(k)$ est symétrique, de polynôme minimal $\mathrm{ppcm}(f_1, \dots, f_r)$, où $n = n_1 + \cdots + n_r$.

Démonstration. Gardons les notations de l'énoncé. Des calculs similaires à ceux faits précédemment montrent que l'on a

$$b_s\big(\ell_x(y_1), y_2\big) = b_s\big(y_1, \ell_x(y_2)\big), \quad \text{pour tous } y_1, y_2 \in E.$$

Le lemme 2.1 montre alors que la matrice M est symétrique. Il reste à calculer le polynôme minimal de ℓ_x. Pour déterminer ce dernier, remarquons que pour tout $P \in k[X]$, on a

$$\begin{aligned} P(\ell_x) = \big(P(\ell_{x_1}), \dots, P(\ell_{x_r})\big) = 0 &\iff P(\ell_{x_j}) = 0, \quad \text{pour tout } j \in [\![1, r]\!] \\ &\iff f_j \,|\, P, \ \text{pour tout } j \in [\![1, r]\!] \\ &\iff \mathrm{ppcm}(f_1, \dots, f_r) \,|\, P, \end{aligned}$$

la deuxième équivalence découlant du fait que le polynôme minimal de ℓ_{x_j} est f_j d'après le lemme 2.2.

Ainsi, les polynômes annulateurs de ℓ_x sont exactement les multiples du ppcm de $f_1, \dots, f_r$. Comme ce dernier est bien unitaire et de degré minimal parmi les polynômes annulateurs non nuls, cela achève la démonstration. □

Finissons ce paragraphe par quelques remarques calculatoires.

2.4. Remarque. Soit k un corps arbitraire. On commence par remarquer le fait suivant. Soit $f = X^n + a_{n-1}X^{n-1} + \cdots + a_1 X + a_0 \in k[X]$ un polynôme unitaire de degré $n \geqslant 1$, et soit $\mathscr{B}_0 = (\overline{1}, \overline{X}, \dots, \overline{X}^{n-1})$ la base canonique de $k[X]/(f)$. Alors, on a $\mathrm{Mat}(\ell_{\overline{X}}\,; \mathscr{B}_0) = C_f$.

En effet, pour tout $j \in [\![0, n-2]\!]$, on a $\ell_{\overline{X}}(\overline{X}^j) = \overline{X}^{j+1}$. De plus, en utilisant la définition de la loi externe sur $k[X]/(f)$, l'égalité $\overline{f} = \overline{0}$ se récrit

$$\overline{X}^n + a_{n-1} \cdot \overline{X}^{n-1} + \cdots + a_0 \cdot \overline{1} = \overline{0},$$

soit $\overline{X}^n = -a_0 \cdot \overline{1} - \cdots - a_{n-1} \cdot \overline{X}^{n-1}$. Par conséquent, on a

$$\mathrm{Mat}(\ell_{\overline{X}}\,;\mathscr{B}_0) = \begin{pmatrix} 0 & 0 & & 0 & -a_0 \\ 1 & 0 & & & -a_1 \\ 0 & 1 & & & -a_2 \\ & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix} = C_f.$$

Reprenons maintenant les notations de la proposition 2.3. Alors, la famille

$$\mathscr{E} = \big((\overline{1},0,\ldots,0),\ldots,(x_1^{n_1-1},0,\ldots,0),\ldots,(0,\ldots,0,\overline{1}),\ldots,(0,\ldots,0,x_r^{n_r-1})\big)$$

est une base de E, et l'on a

$$\mathrm{Mat}(\ell_x\,;\mathscr{E}) = \begin{pmatrix} C_{f_1} & & \\ & \ddots & \\ & & C_{f_r} \end{pmatrix}.$$

Si P est la matrice des vecteurs de $\mathscr{B}$ dans la base $\mathscr{E}$, une matrice symétrique de $\mathrm{M}_n(k)$ de polynôme minimal $\mathrm{ppcm}(f_1,\ldots,f_r)$ est donc

$$M = \mathrm{Mat}(\ell_x\,;\mathscr{B}) = P^{-1} \begin{pmatrix} C_{f_1} & & \\ & \ddots & \\ & & C_{f_r} \end{pmatrix} P.$$

Dans ce qui suit, nous appliquerons cette méthode pour démontrer les résultats annoncés dans l'introduction.

Nous commencerons par le cas où $-1 \in k^{\times 2}$, qui est un peu moins technique.

3. Le cas $-1 \in k^{\times 2}$

Les résultats que nous avons en vue dans ce paragraphe sont les suivants.

3.1. Théorème. *Soit k un corps de caractéristique différente de 2. On suppose que $-1 \in k^{\times 2}$.*

Alors, tout polynôme unitaire de $k[X]$ de degré $n \geqslant 1$ est le polynôme minimal d'une matrice de $\mathscr{S}_n(k)$.

3.2. Corollaire. *Soit k un corps de caractéristique différente de 2. On suppose que $-1 \in k^{\times 2}$. Alors, tout élément de $\overline{k}$ de degré n sur k est valeur propre d'une matrice de $\mathscr{S}_n(k)$.*

Afin de pouvoir appliquer la proposition 2.3, nous aurons besoin de résultats d'algèbre bilinéaire.

3.3. Proposition. *Soit $b : V \times V \longrightarrow k$ une forme bilinéaire symétrique, et soit U un sous-espace vectoriel de V de dimension finie tel que la restriction de b à $U \times U$ est non dégénérée. Alors, on a $V = U \oplus U^{\perp}$.*

Si de plus b est non dégénérée, alors la restriction de b à $U^{\perp} \times U^{\perp}$ est aussi non dégénérée.

Démonstration. Soit $x \in V$, et soit $\varphi = b(x,\_)_{|U} \in U^*$. Puisque U est de dimension finie et la restriction de b à $U \times U$ est non dégénérée, l'application

$$\begin{aligned} U &\longrightarrow U^* \\ u &\longmapsto b(u,\_) \end{aligned}$$

est un isomorphisme. La forme linéaire φ a donc un antécédent, ce qui revient à dire qu'il existe $u \in U$ tel que $b(u,u') = b(x,u')$ pour tout $u' \in U$. Par conséquent, $b(x-u,u') = 0$ pour tout $u' \in U$, soit $x - u \in U^{\perp}$. On en déduit que $x = u + (x-u) \in U + U^{\perp}$. De plus, $U \cap U^{\perp} = \{0\}$. En effet, si $x \in U \cap U^{\perp}$, alors $b(x,u') = 0$ pour tout $u' \in U$. Comme $x \in U$, la non-dégénérescence de la restriction de b à $U \times U$ entraîne que $x = 0$. On a donc bien $V = U \oplus U^{\perp}$.

Supposons maintenant que b soit également non dégénérée. Soit $x \in U^{\perp}$ tel que $b(x,y') = 0$ pour $y' \in U^{\perp}$. Alors, $b(x,y) = 0$ pour tout $y \in V$. En effet, d'après ce qui précède, tout $y \in V$ s'écrit sous la forme $y = y_U + y_{U^{\perp}}$, avec $y_U \in U$ et $y_{U^{\perp}} \in U^{\perp}$. Mais alors, on a $b(x,y_U) = 0$ puisque $x \in U^{\perp}$ et $y_U \in U$, et $b(x,y_{U^{\perp}}) = 0$ par hypothèse sur x. Ainsi, $b(x,y) = 0$ pour tout $y \in V$. Puisque b est non dégénérée, on obtient finalement $x = 0$. Cela achève la démonstration. □

On peut maintenant énoncer et démontrer le théorème suivant.

3.4. Théorème. *Soit k un corps de caractéristique différente de 2.*

Soit $b : V \times V \longrightarrow k$ une forme bilinéaire symétrique non dégénérée, où V est de dimension $\dim_k(V) = 2m \geqslant 2$. On suppose qu'il existe un sous-espace W de dimension m tel que $b(x,y) = 0$ pour tous $x, y \in W$.

Alors, il existe une base $\mathscr{B}$ de V telle que

$$\mathrm{Mat}(b\,;\mathscr{B}) = \begin{pmatrix} 1 & & & & \\ & -1 & & & \\ & & \ddots & & \\ & & & 1 & \\ & & & & -1 \end{pmatrix}.$$

En particulier, si $-1 \in k^{\times 2}$, l'espace V possède une base b-orthonormée.

Démonstration. On se fixe une base $(f_1, \dots, f_m)$ de W. Soit $\varphi \in V^*$ une forme linéaire telle que

$$\varphi(f_1) = 1, \quad \varphi(f_2) = \cdots = \varphi(f_m) = 0.$$

Une telle forme linéaire existe : il suffit de compléter $(f_1, \dots, f_m)$ en une base de V, et de choisir arbitrairement l'image des autres vecteurs de base.

Puisque b est non dégénérée, il existe $v \in V$ tel que $b(v, \_) = \varphi$. On a donc

$$b(v, f_1) = 1, \quad b(v, f_2) = \cdots = b(v, f_m) = 0.$$

Posons $e_1 = v - \dfrac{b(v,v)}{2} f_1$. Comme $b(f_i, f_j) = 0$ pour tous $i, j \in [\![1, m]\!]$, on obtient aisément

$$b(e_1, f_1) = 1, \quad b(e_1, f_2) = \cdots = b(e_1, f_m) = 0.$$

De plus, on a

$$b(e_1, e_1) = b(v,v) - 2\frac{b(v,v)}{2} b(v, f_1) + \Big(\frac{b(v,v)}{2}\Big)^2 b(f_1, f_1) = 0.$$

Notons maintenant que la famille (e_1, f_1) est k-libre. En effet, soient $\lambda, \mu \in k$ tels que $\lambda e_1 + \mu f_1 = 0$. Alors,

$$b(\lambda e_1 + \mu f_1, f_1) = 0 = \lambda b(e_1, f_1) + \mu b(f_1, f_1) = \lambda.$$

Puisque f_1 est non nul, on en déduit alors $\mu = 0$.

Soit $U = \mathrm{Vect}_k(e_1, f_1)$. La restriction de b à $U \times U$ est non dégénérée, puisque la matrice de cette restriction dans la base (e_1, f_1) est $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, qui est de déterminant -1. Par conséquent, on a l'égalité $V = U \oplus U^\perp$, et la restriction de b à $U^\perp \times U^\perp$ est non dégénérée, d'après la proposition précédente.

De plus, $U^\perp$ est un sous-espace de dimension $2m - 2$, et contient un sous-espace W' de dimension $m - 1$ tel que $b(x, y) = 0$ pour tous $x, y \in W'$. En effet, par hypothèse sur W, on a $b(f_1, f_2) = \cdots = b(f_1, f_m) = 0$, et par construction de e_1, on a $b(e_1, f_2) = \cdots = b(e_1, f_m) = 0$. Par conséquent, on obtient $f_2, \dots, f_m \in U^\perp$, et par suite $W' = \mathrm{Vect}_k(f_2, \dots, f_m) \subset U^\perp$.

Remarquons enfin que si l'on pose $u_1 = e_1 + \dfrac{1}{2} f_1$ et $v_1 = e_1 - \dfrac{1}{2} f_1$, alors la famille (u_1, v_1) est une base de U vérifiant

$$b(u_1, u_1) = 1, \quad b(u_1, v_1) = 0 \quad \text{et} \quad b(v_1, v_1) = -1.$$

Une récurrence immédiate permet alors de conclure quant à l'existence d'une base $\mathscr{B}$ vérifiant les propriétés de l'énoncé.

Supposons maintenant que $-1 \in k^\times$. Soit $i \in k$ tel que $i^2 = -1$, et choisissons une base b-orthogonale $\mathscr{B} = (u_1, v_1, \ldots, u_m, v_m)$ telle que $b(u_j, u_j) = 1$ et $b(v_j, v_j) = -1$ pour tout $j \in [\![1, m]\!]$. Il est alors facile de voir que la famille $(u_1, iv_1, \ldots, u_m, iv_m)$ est une base b-orthonormée. □

D'après la proposition 2.3, le théorème 3.1 découle immédiatement du résultat suivant.

3.5. Proposition. *Soit k un corps de caractéristique différente de 2. On suppose que $-1 \in k^{\times 2}$. Soit $f \in k[X]$ un polynôme unitaire de degré $n \geqslant 1$, et soit $s : k[X]/(f) \longrightarrow k$ l'unique forme linéaire définie par*

$$s(\overline{1}) = s(x) = \cdots = s(x^{n-2}) = 0, \quad s(x^{n-1}) = 1,$$

où $x = \overline{X}$. Alors, $k[X]/(f)$ possède une base b_s-orthonormée.

Démonstration. Commençons par démontrer que b_s est non dégénérée. Soit $y = \sum_{j=0}^{n-1} \lambda_j \cdot x^j \in k[X]/(f)$ tel que $s(yz) = 0$ pour tout $z \in k[X]/(f)$. En particulier, on a

$$s(y) = \lambda_{n-1} = 0.$$

Supposons avoir démontré que $\lambda_{n-1} = \cdots = \lambda_{n-1-\ell} = 0$ pour un certain entier $\ell \in [\![0, n-2]\!]$. On a donc $y = \sum_{j=0}^{n-\ell-2} \lambda_j \cdot x^j$. Par suite,

$$s(yx^{\ell+1}) = s\Big(\sum_{j=0}^{n-\ell-2} \lambda_j \cdot x^{j+\ell+1}\Big) = s\Big(\sum_{j=\ell+1}^{n-1} \lambda_{j-(\ell+1)} \cdot x^j\Big) = \lambda_{n-1-(\ell+1)}.$$

Une récurrence immédiate montre alors que $\lambda_{n-1} = \cdots = \lambda_0 = 0$, c'est-à-dire $y = 0$.

Montrons maintenant l'existence d'une base b_s-orthonormée. Supposons d'abord que n soit impair. Écrivons $n = 2m + 1$, et posons $V = kx^m$. Puisque $b_s(x^m, x^m) = s(x^{2m}) = s(x^{n-1}) = 1$, la restriction de b_s à V est non dégénérée. D'après la proposition 3.3, on a donc $k[X]/(f) = V \oplus V^\perp$. Comme $b_s(x^m, x^m) = 1, x^m$ est une base b_s-orthonormée de V. Il reste donc à montrer l'existence d'une base b_s-orthonormée de $V^\perp$ pour conclure.

Puisque b_s est non dégénérée, la proposition 3.3 montre que la restriction de b_s à $V^\perp$ est également non dégénérée. Soit $W = \mathrm{Vect}(1, x, \ldots, x^{m-1})$. Alors, W est un sous-espace de $V^\perp$ de dimension m. En effet, pour tout entier $j \in [\![0, m-1]\!]$, on a $0 \leqslant j + m \leqslant 2m - 1 < n - 1$, d'où

$$b_s(x^j, x^m) = s(x^{j+m}) = 0.$$

Enfin, pour tous j, $\ell \in [\![0, m-1]\!]$, on a $0 \leqslant j+\ell \leqslant 2m-2 < n-1$, et ainsi

$$b_s(x^j, x^\ell) = s(x^{j+\ell}) = 0.$$

Par bilinéarité, on en déduit que $b_s(y,z) = 0$ pour tous $y, z \in W$. Mais, le sous-espace $V^\perp$ est de dimension $2m$, et le théorème 3.4 nous donne alors l'existence d'une base b_s-orthonormée de $V^\perp$. Cela achève la démonstration dans le cas impair.

Supposons maintenant que n soit pair, et écrivons $n = 2m$, avec $m \geqslant 1$. Soit $W = \mathrm{Vect}_k(1, x, \ldots, x^{m-1})$. Alors, pour tous entiers $j, \ell \in [\![0, m-1]\!]$, on a $0 \leqslant j+\ell \leqslant 2m-2 < n-1$, et ainsi

$$b_s(x^j, x^\ell) = s(x^{j+\ell}) = 0.$$

Par bilinéarité, on en déduit que $b_s(y,z) = 0$ pour tous $y, z \in W$. Comme le sous-espace W est de dimension m, le théorème 3.4 entraîne alors l'existence d'une base b_s-orthonormée de $k[X]/(f)$. □

Le corollaire 3.2 se déduit alors immédiatement du théorème 3.1, appliqué au polynôme $\mu_{\alpha,k}$.

Nous nous proposons maintenant d'appliquer ce qui précède à $f = X^3 - d$, où $d \in k$. Si $s : k[X]/(f) \longrightarrow k$ est définie comme dans la proposition 3.5 et si $x = \overline{X}$, on a $s(\overline{1}) = s(x) = 0, s(x^2) = 1$ et aussi

$$s(x^3) = s(\overline{d}) = s(d\cdot\overline{1}) = d\cdot s(\overline{1}) = 0, \quad s(x^4) = s(d\cdot x) = d\cdot s(x) = 0.$$

On a donc

$$\mathrm{Mat}(b_s\,;\mathscr{E}) = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix},$$

où $\mathscr{E} = (\overline{1}, x, x^2)$.

On vérifie alors que la famille $\mathscr{B} = \big(x, \overline{1} + \frac{1}{2}\cdot x^2,\ i\cdot(\overline{1} - \frac{1}{2}x^2)\big)$, où $i \in k$ satisfait $i^2 = -1$, est une base b_s-orthonormée. D'après la remarque 2.4, si $P = \begin{pmatrix} 0 & 1 & i \\ 1 & 0 & 0 \\ 0 & \frac{1}{2} & -\frac{i}{2} \end{pmatrix}$, la matrice

$$M = \mathrm{Mat}(\ell_x\,;\mathscr{B}) = P^{-1}C_f P = \begin{pmatrix} 0 & 1 & i \\ 1 & \frac{d}{4} & -\frac{id}{4} \\ i & -\frac{id}{4} & -\frac{d}{4} \end{pmatrix}$$

est symétrique de polynôme minimal $X^3 - d$.

On retrouve ainsi la matrice donnée dans l'introduction.

4. Le cas $k = \mathbb{Q}$

Nous commençons par énoncer les résultats principaux que nous démontrerons dans ce paragraphe (cf. [14]).

4.1. Théorème. *Soit $f \in \mathbb{Q}[X]$ un polynôme unitaire de degré $n \geqslant 1$ scindé sur $\mathbb{R}$ sans racines multiples. Alors, f est le polynôme minimal d'une matrice de $\mathscr{S}_{4n}(\mathbb{Q})$.*

4.2. Corollaire. *Soit α un nombre totalement réel de degré $n \geqslant 1$ sur $\mathbb{Q}$. Alors, α est valeur propre d'une matrice de $\mathscr{S}_{4n}(\mathbb{Q})$.*

4.3. Remarque. Notons que c'est Krakowski qui a démontré le premier qu'un nombre totalement réel est valeur propre d'une matrice symétrique à coefficients rationnels (cf. [32]). Néanmoins, la taille de la matrice en question était exponentielle par rapport au degré du nombre considéré. Le corollaire précédent est donc bien meilleur, et l'on peut faire encore mieux, comme on le verra plus loin.

La démonstration du théorème 4.1 s'appuiera sur le théorème des quatre carrés, que l'on énonce seulement dans le cas rationnel, et dont nous reléguons la démonstration à un paragraphe ultérieur.

4.4. Théorème. *Tout entier rationnel positif est somme de quatre carrés de rationnels.*

Nous allons en déduire le résultat d'algèbre bilinéaire suivant.

4.5. Théorème. *Soit $b : V \times V \longrightarrow \mathbb{Q}$ une forme bilinéaire symétrique, où V est un $\mathbb{Q}$-espace vectoriel de dimension finie.*

On suppose que $b(x,x) > 0$ pour tout $x \in V$ non nul. Alors, V^4 possède une base orthonormée pour la forme bilinéaire symétrique

$$\begin{aligned} \theta_b \ : \quad V^4 \times V^4 &\longrightarrow \mathbb{Q} \\ \big((x_1,x_2,x_3,x_4),(y_1,y_2,y_3,y_4)\big) &\longmapsto \sum_{i=1}^{4} b(x_i,y_i). \end{aligned}$$

Démonstration. Soit $(e_1,\ldots,e_n)$ une base b-orthogonale de V fixée. Pour tout entier $i \in [\![1,n]\!]$, posons $r_i = b(e_i,e_i)$. Il est alors facile de voir que la famille

$$\mathscr{B} = \big((e_1,0,0,0),(0,e_1,0,0),(0,0,e_1,0),(0,0,0,e_1),\ldots,(0,0,0,e_n)\big)$$

est une base θ-orthogonale de V^4, et que la matrice représentative de θ_b

dans cette base est

$$M = \begin{pmatrix} r_1 \mathrm{I}_4 & & \\ & \ddots & \\ & & r_n \mathrm{I}_4 \end{pmatrix}.$$

Nous allons exhiber une matrice $P \in \mathrm{GL}_{4n}(\mathbb{Q})$ telle que $P^t M P = \mathrm{I}_{4n}$. Cela suffira à démontrer le théorème. En effet, si $\mathscr{B}' = (f_1, \ldots, f_{4n})$ est la base dont la matrice des vecteurs coordonnées dans la base $\mathscr{B}$ est P, alors la formule de changement de base pour les matrices représentatives de formes bilinéaires donne l'égalité $\mathrm{Mat}(\theta_b\,; \mathscr{B}') = P^t M P = \mathrm{I}_{4n}$, ce qui démontre que $\mathscr{B}'$ est θ_b-orthonormée.

Il suffit en fait de montrer que pour tout rationnel $r > 0$, il existe une matrice $Q \in \mathrm{GL}_4(\mathbb{Q})$ telle que $rQ^tQ = \mathrm{I}_4$. En effet, par hypothèse sur b, chaque r_i est > 0. Si $Q_i \in \mathrm{GL}_4(\mathbb{Q})$ vérifie $r_i Q_i^t Q_i = \mathrm{I}_4$, il est alors aisé de voir que la matrice $P = \begin{pmatrix} Q_1 & & \\ & \ddots & \\ & & Q_n \end{pmatrix}$ convient.

Soit donc $r > 0$ un rationnel strictement positif. D'après le théorème 4.4, il existe $a, b, c, d \in \mathbb{Q}$ tels que $r = a^2 + b^2 + c^2 + d^2$.

Supposons tout d'abord que $c^2 + d^2 = 0$, i.e. $r = a^2 + b^2$. On vérifie alors que la matrice

$$Q = \frac{1}{r} \begin{pmatrix} a & b & 0 & 0 \\ b & -a & 0 & 0 \\ 0 & 0 & a & b \\ 0 & 0 & b & -a \end{pmatrix}$$

satisfait $rQ^tQ = \mathrm{I}_4$. Supposons maintenant que $c^2 + d^2 \neq 0$. Cette fois, on vérifie que la matrice

$$Q = \frac{1}{r} \begin{pmatrix} a & b & c & d \\ b & -a & d & -c \\ c & d & \dfrac{-ac^2 + ad^2 - 2bcd}{c^2 + d^2} & \dfrac{-2acd + bc^2 - bd^2}{c^2 + d^2} \\ d & -c & \dfrac{-2acd + bc^2 - bd^2}{c^2 + d^2} & \dfrac{ac^2 - ad^2 + 2bcd}{c^2 + d^2} \end{pmatrix}$$

satisfait $rQ^tQ = \mathrm{I}_4$. Cela achève la démonstration. □

L'idée est donc d'appliquer le théorème 4.5 à une forme bilinéaire b bien choisie. Supposons un instant que nous ayons à disposition une forme linéaire $t : \mathbb{Q}[X]/(f) \longrightarrow \mathbb{Q}$ telle que l'on ait

$$b_t(y, y) = t(y^2) > 0, \quad \text{pour tout } y \in \mathbb{Q}[X]/(f) \text{ non nul},$$

et soit $s : \big(\mathbb{Q}[X]/(f)\big)^4 \longrightarrow \mathbb{Q}$ la forme linéaire définie par

$$s\big((y_1, y_2, y_3, y_4)\big) = \sum_{i=1}^{4} t(y_i).$$

Alors, la forme bilinéaire b_s est exactement la forme bilinéaire θ_{b_t}. En effet, pour tous $(y_1, y_2, y_3, y_4), (z_1, z_2, z_3, z_4) \in \big(\mathbb{Q}[X]/(f)\big)^4$, on a

$$\begin{aligned} b_s\big((y_1, y_2, y_3, y_4), (z_1, z_2, z_3, z_4)\big) &= s\big((y_1, y_2, y_3, y_4)(z_1, z_2, z_3, z_4)\big) \\ &= s\big((y_1 z_1, y_2 z_2, y_3 z_3, y_4, z_4)\big) \\ &= \sum_{i=1}^{4} t(y_i z_i) \\ &= \sum_{i=1}^{4} b_t(y_i, z_i) \\ &= \theta_{b_t}\big((y_1, y_2, y_3, y_4), (z_1, z_2, z_3, z_4)\big). \end{aligned}$$

Mais alors, le théorème 4.5 montre que la $\mathbb{Q}$-algèbre $\big(\mathbb{Q}[X]/(f)\big)^4$ possède une base b_s-orthonormée. Une application de la proposition 2.3 montre alors que f est le polynôme minimal d'une matrice symétrique de $\mathrm{M}_{4n}(\mathbb{Q})$. Il reste donc à exhiber une telle forme linéaire $t : \mathbb{Q}[X]/(f) \longrightarrow \mathbb{Q}$, ce qui est le but du lemme suivant.

4.6. Lemme. *Soit $f \in \mathbb{Q}[X]$ un polynôme unitaire de degré $n \geqslant 1$ scindé sur $\mathbb{R}$ sans racines multiples, et soit $E = \mathbb{Q}[X]/(f)$.*
Alors, on a $\mathrm{Tr}_{E/\mathbb{Q}}(y^2) > 0$ pour tout $y \in E$ non nul.

Démonstration. Soit $x = \overline{X}$. Puisque $\mathscr{B}_0 = (\overline{1}, x, \ldots, x^{n-1})$ est une $\mathbb{Q}$-base de E, un élément $y \in E$ s'écrit sous la forme $y = \sum_{k=0}^{n-1} a_i \cdot x^k$, $a_i \in k$. On a alors $\ell_y = P(\ell_x)$, où $P = \sum_{k=0}^{n-1} a_k X^k$, puis $\ell_{y^2} = \ell_y^2 = \big(P(\ell_x)\big)^2$. Par conséquent, on a

$$\mathrm{Mat}(\ell_{y^2}\,; \mathscr{B}_0) = \big(P(C_f)\big)^2.$$

Comme f est le polynôme minimal de ℓ_x d'après le lemme 2.2, et que f est scindé sur $\mathbb{R}$ sans racines multiples, $\mathrm{Mat}(\ell_x\,; \mathscr{B}_0) = C_f$ est diagonalisable sur $\mathbb{R}$. Il existe donc $Q \in \mathrm{GL}_n(\mathbb{R})$ telle que

$$C_f = Q^{-1} \begin{pmatrix} \alpha_1 & & \\ & \ddots & \\ & & \alpha_n \end{pmatrix} Q,$$

où $\alpha_1, \ldots, \alpha_n \in \mathbb{R}$ sont les racines de f. Mais alors, on a

$$\mathrm{Mat}(\ell_{y^2}\,;\mathscr{B}_0) = Q^{-1}\begin{pmatrix} P(\alpha_1)^2 & & \\ & \ddots & \\ & & P(\alpha_n)^2 \end{pmatrix} Q,$$

et par conséquent,

$$\mathrm{Tr}_{E/\mathbb{Q}}(y^2) = \sum_{j=1}^{n} P(\alpha_i)^2 \geqslant 0.$$

Soit maintenant $y \in E$ tel que $\mathrm{Tr}_{E/\mathbb{Q}}(y^2) = 0$, et montrons que $y = 0$. En gardant les notations précédentes, on a donc

$$P(\alpha_1) = \cdots = P(\alpha_n) = 0.$$

Comme $\alpha_1, \ldots, \alpha_n$ sont des réels deux à deux distincts par hypothèse, P est un polynôme de degré $\leqslant n-1$ possédant au moins n racines distinctes. Par conséquent, $P = 0$, et donc $y = 0$. □

Nous avons donc établi le théorème 4.1.

Le corollaire 4.2 se déduit alors du théorème 4.1, appliqué à $\mu_{\alpha,\mathbb{Q}}$, qui est un polynôme irréductible, donc séparable puisque $\mathbb{Q}$ est de caractéristique nulle.

Ce résultat est très loin d'être optimal. Si $d \in \mathbb{Q}$ est une somme de deux carrés de rationnels, mais pas le carré d'un rationnel, alors $\sqrt{d}$ est totalement réel de degré 2 sur $\mathbb{Q}$. Le corollaire 4.2 montre alors que $\sqrt{d}$ est valeur propre d'une matrice de $\mathscr{S}_8(\mathbb{Q})$. Néanmoins, comme d est somme de deux carrés, la proposition 1.4 montre que $\sqrt{d}$ est valeur propre d'une matrice de $\mathscr{S}_2(\mathbb{Q})$.

Le même phénomène se produit pour $\sqrt{3}$. Le corollaire 4.2 fournit l'existence d'une matrice de $\mathscr{S}_8(\mathbb{Q})$ dont $\sqrt{3}$ est valeur propre, alors que l'introduction fournit une matrice de $\mathscr{S}_3(\mathbb{Q})$ dont $\sqrt{3}$ est valeur propre. Notons au passage que ce dernier résultat est optimal pour $\sqrt{3}$, puisque $\sqrt{3}$ n'est pas valeur propre d'une matrice de $\mathscr{S}_2(\mathbb{Q})$.

En fait, nous avons le théorème quasi optimal suivant, dont la démonstration est extrêmement difficile et utilise des outils très sophistiqués (cf. [13]).

4.7. Théorème (Bender). *Soit α un nombre totalement réel, et soit $n \geqslant 1$ son degré sur $\mathbb{Q}$.*

(1) *Si n est impair, α est valeur propre d'une matrice de $\mathscr{S}_n(\mathbb{Q})$.*

(2) *Si n est pair, α est valeur propre d'une matrice de $\mathscr{S}_{n+1}(\mathbb{Q})$.*

4.8. Remarque. À l'heure actuelle, on ne sait pas caractériser les nombres totalement réels de degré n pair qui ne sont pas valeur propre d'une matrice de $\mathscr{S}_n(\mathbb{Q})$.

Nous nous proposons maintenant de retrouver les résultats de l'introduction en utilisant le point de vue exposé dans ce chapitre.

Soit $d = u^2 + v^2 \in \mathbb{Q}$. On suppose que d n'est pas un carré, si bien que l'on peut supposer que $v \neq 0$. Soit $f = X^2 - d$, et soit $s : \mathbb{Q}[X]/(f) \longrightarrow \mathbb{Q}$ l'unique forme linéaire telle que

$$s(\overline{1}) = 1 \quad \text{et} \quad s(x) = u.$$

Il est alors facile de constater que la famille $\mathscr{B} = \big(\overline{1}, v^{-1}\cdot(u\cdot\overline{1} - x)\big)$ est une base b_s-orthonormée de $\mathbb{Q}[X]/(f)$. La matrice de changement de base correspondante est donc $P = \begin{pmatrix} 1 & v^{-1}u \\ 0 & -v^{-1} \end{pmatrix}$. La remarque 2.4 montre alors que

$$\mathrm{Mat}(\ell_x\,;\mathscr{B}) = P^{-1}C_fP = \begin{pmatrix} u & v^{-1}(u^2-d) \\ -v & -u \end{pmatrix},$$

et puisque $d = u^2 + v^2$, on a $\mathrm{Mat}(\ell_x\,;\mathscr{B}) = \begin{pmatrix} u & -v \\ -v & -u \end{pmatrix}$, qui est la matrice donnée dans l'introduction.

Nous allons maintenant expliquer comment nous avons obtenu la matrice carrée 3×3 symétrique de polynôme caractéristique égal à $X(X^2 - 3)$. Soit $f = X^2 - 3$. Si l'on note $t : \mathbb{Q}[X]/(f) \longrightarrow \mathbb{Q}$ la forme linéaire du lemme 4.6, de simples calculs montrent que l'on a

$$\mathrm{Mat}(b_t\,;\mathscr{B}_0) = \begin{pmatrix} 2 & 0 \\ 0 & 6 \end{pmatrix}.$$

Notons maintenant que, si $b : V \times V \longrightarrow k$ est une forme bilinéaire symétrique, et si V possède une une base b-orthonormée, alors le déterminant de n'importe quelle matrice représentative de b est un carré d'un élément non nul de k. En effet, si $\mathscr{B}$ est une base b-orthonormée, on a $\mathrm{Mat}(b\,;\mathscr{B}) = \mathrm{I}_n$. Si maintenant $\mathscr{B}'$ est une base arbitraire de V, et si P est la matrice des vecteurs de $\mathscr{B}'$ dans la base $\mathscr{B}$, on a $\mathrm{Mat}(b\,;\mathscr{B}') = P^tP$, qui est de déterminant égal à $\det(P)^2$.

L'idée sous-jacente est donc de compléter la matrice précédente en une matrice symétrique 3×3 dont le déterminant est un carré, et espérer que la forme bilinéaire symétrique qu'elle représente admette une base orthonormée. Pour concrétiser cette idée fort vague, nous allons utiliser la proposition 2.3, avec $r = 2$, $f_1 = X$ et $f_2 = f = X^2 - 3$. Notons que $\mathbb{Q}[X]/(f_1)$ n'est rien d'autre que $\mathbb{Q}$ (car c'est un $\mathbb{Q}$-espace vectoriel de dimension 1).

Pour appliquer cette dernière proposition, il faut donc trouver une forme linéaire $s : \mathbb{Q} \times \mathbb{Q}[X]/(f) \longrightarrow \mathbb{Q}$ appropriée. Afin de poursuivre l'idée précédente, il semble donc tout indiqué de poser

$$s\big((r,y)\big) = 3r + t(y), \quad \text{pour tout } (r,y) \in \mathbb{Q} \times \mathbb{Q}[X]/(f).$$

En effet, si l'on prend la base $\mathscr{E}$ de la remarque 2.4, on voit facilement que l'on a

$$B = \mathrm{Mat}(b_s\,;\mathscr{E}) = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 6 \end{pmatrix},$$

qui est de déterminant 6^2.

Pour alléger un peu la rédaction, nous allons maintenant introduire une notation commode.

Si k est un corps, $M_1, M_2 \in \mathrm{M}_n(k)$, et si $P \in \mathrm{GL}_n(k)$, on note $M_1 \underset{P}{\sim} M_2$ lorsque $P^t M_1 P = M_2$.

Afin de pouvoir faire les calculs, on va utiliser l'identité suivante. Si $d \in \mathbb{Q}$, et si $r = a^2 + db^2$, on a

$$\begin{pmatrix} r & 0 \\ 0 & rd \end{pmatrix} \underset{Q}{\sim} \begin{pmatrix} 1 & 0 \\ 0 & d \end{pmatrix}, \quad \text{avec } Q = r^{-1}\begin{pmatrix} a & db \\ b & -a \end{pmatrix}.$$

On a alors, successivement,

$$(1)\ \ B \underset{P_1}{\sim} B_2 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 6 \end{pmatrix}, \quad \text{avec } P_1 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix};$$

$$(2)\ \ B_2 \underset{P_2}{\sim} B_3 = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad \text{avec } P_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{3} & \frac{2}{3} \\ 0 & \frac{1}{3} & -\frac{1}{3} \end{pmatrix};$$

$$(3)\ \ B_3 \underset{P_1}{\sim} B_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix};$$

$$(4)\ \ B_4 \underset{P_3}{\sim} \mathrm{I}_3, \quad \text{avec } P_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & \frac{1}{2} & -\frac{1}{2} \end{pmatrix}.$$

En recollant les morceaux, on a donc $P^t M P = \mathrm{I}_3$, avec

$$P = P_1 P_2 P_1 P_3 = \frac{1}{6}\begin{pmatrix} 2 & 2 & -2 \\ 0 & 3 & 3 \\ 2 & -1 & 1 \end{pmatrix}.$$

La matrice P est alors la matrice des vecteurs coordonnées d'une base $\mathscr{B}$ dans la base $\mathscr{E}$, et c'est la base b_s-orthonormée recherchée.

La remarque 2.4 montre alors qu'une matrice symétrique de polynôme minimal $\operatorname{ppcm}(X, X^2-3) = X(X^2-3)$ est donnée par

$$M = P^{-1}\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 3 \\ 0 & 1 & 0 \end{pmatrix} P = \begin{pmatrix} 0 & 1 & 1 \\ 1 & -1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

4.9. Remarque. Les diverses démonstrations et calculs qui précèdent peuvent paraître un peu miraculeux. En particulier, le théorème 4.5 et sa démonstration peuvent sembler un peu sortis de nulle part aux yeux d'un lecteur non spécialiste d'algèbre bilinéaire, et ce, à juste titre.

Cet état de fait découle directement du parti pris choisi, à savoir essayer de limiter l'introduction de nouveaux concepts au strict minimum. Néanmoins, pour le lecteur curieux de comprendre un peu plus en profondeur, nous révélons tout ce qui a été caché soigneusement sous le tapis dans le paragraphe 6.

5. Démonstration du théorème des quatre carrés

Nous démontrons ici la version rationnelle du théorème des quatre carrés de Lagrange (cf. théorème 4.4), à savoir que tout rationnel positif est somme de quatre carrés de rationnels.

Notons tout d'abord qu'il suffit de démontrer ce résultat dans le cas des entiers positifs. En effet, supposons que cela soit le cas et soit $x \in \mathbb{Q}$ un rationnel positif. On peut alors toujours écrire $x = \dfrac{a}{b}$, où a, b sont des entiers positifs et b est non nul. Mais alors, $ab = u_1^2+u_2^2+u_3^2+u_4^2$, $u_i \in \mathbb{Q}$, et l'on a

$$x = \frac{ab}{b^2} = \frac{u_1^2+u_2^2+u_3^2+u_4^2}{b^2} = x_1^2+x_2^2+x_3^2+x_4^2,$$

avec $x_i = b^{-1}u_i$.

Il suffit donc de démontrer que tout entier $n \geqslant 0$ est somme de quatre carrés de rationnels. Les cas de 0 et 1 étant triviaux, on peut toujours supposer que $n \geqslant 2$.

Nous utiliserons, à plusieurs reprises, l'identité suivante (due à Euler), un peu effrayante au fond [3], mais qui se démontre fort aisément en développant

3. Ce qui nous fait dire que le fond d'Euler effraie...

tout simplement :

$$\begin{aligned}(x_1^2+y_1^2+z_1^2+t_1^2)(x_2^2+y_2^2+z_2^2+t_2^2) &= (x_1x_2+y_1y_2+z_1z_2+t_1t_2)^2 \\ &\quad +(x_1y_2-y_1x_2+t_1z_2-z_1t_2)^2 \\ &\quad +(x_1z_2-z_1x_2+y_1t_2-t_1y_2)^2 \\ &\quad +(x_1t_2-t_1x_2+z_1y_2-y_1z_2)^2.\end{aligned}$$

pour tous $x_1, x_2, y_1, y_2, z_1, z_2, t_1, t_2 \in \mathbb{Q}$.

Autrement dit, un produit de deux sommes de quatre carrés de rationnels est encore une somme de quatre carrés de rationnels.

On suppose maintenant que l'ensemble des entiers $n \geqslant 2$ qui ne s'écrivent pas comme somme de quatre carrés de rationnels est non vide, et soit p le plus petit élément de cet ensemble. Alors, p est premier. Dans le cas contraire, on aurait $p = n_1 n_2$, avec $2 \leqslant n_1, n_2 \leqslant p-1$. Par minimalité de p, les entiers n_1 et n_2 sont des sommes de quatre carrés de rationnels, et il est en alors de même de p d'après l'identité d'Euler. Remarquons aussi que p est impair, puisque $2 = 1^2 + 1^2 + 0^2 + 0^2$.

On commence par établir le fait suivant.

Fait. Pour tout nombre premier impair p, il existe des entiers $a, b \geqslant 0$ et un entier $n \in [\![1, p-1]\!]$ tels que $np = 1 + a^2 + b^2$.

Démonstration du fait. Remarquons que pour tous $x, y \in \left[\!\left[0, \frac{p-1}{2}\right]\!\right]$, on a

$$x^2 \equiv y^2 \ [p] \implies x = y.$$

En effet, si $x^2 \equiv y^2$ $[p]$, alors $(x-y)(x+y) \equiv 0$ $[p]$, soit $p \mid (x-y)(x+y)$. Comme p est premier, $p \mid (x-y)$ ou $p \mid (x+y)$. Mais, on a

$$|x \pm y| \leqslant |x| + |y| \leqslant p-1 < p.$$

Par conséquent, $x \pm y$ est un multiple de p strictement compris entre $-p$ et p, c'est-à-dire 0. On a donc soit $x = y$, soit $x = -y$. Comme $x, y \geqslant 0$, le second cas implique $x = 0 = y$. Bref, dans les deux cas, $x = y$.

Les $\frac{p+1}{2}$ entiers a^2, $a \in \left[\!\left[0, \frac{p-1}{2}\right]\!\right]$, sont donc deux à deux non congrus modulo p. De même, les $\frac{p+1}{2}$ entiers $-b^2 - 1$, $b \in \left[\!\left[0, \frac{p-1}{2}\right]\!\right]$, sont deux à deux non congrus modulo p. Par le principe des tiroirs, il existe donc deux entiers $a, b \in \left[\!\left[0, \frac{p-1}{2}\right]\!\right]$ tels que a^2 et $-b^2 - 1$ sont congrus modulo p, c'est-à-dire pour lesquels $a^2 + b^2 + 1 = np$, avec $n \in \mathbb{Z}$. Notons que $1 \leqslant np \leqslant \frac{(p-1)^2}{2} + 1 < p^2$, et donc que $n \in [\![1, p-1]\!]$.

Passons maintenant à la démonstration proprement dite. Le fait précédent montre en particulier l'ensemble des entiers $n \in [\![1, p-1]\!]$ tels que np est somme de quatre carrés d'entiers est non vide. Si $n = 1$, on a une contradiction, puisque p n'est pas somme de quatre carrés de rationnels. Ainsi, on a $n \geqslant 2$. Comme $n < p$, n est somme de quatre carrés de rationnels par minimalité de p. Mais alors, $n^2 p = n(np)$ est somme de quatre carrés de rationnels d'après l'identité d'Euler, et il en est de même de $p = (n^2)^{-1}(n^2 p)$, d'où encore une contradiction avec le choix de p. Cela achève la démonstration. □

6. Tout ce qui vous a été caché...

Le but de ce paragraphe est de donner un point de vue plus conceptuel, mais plus éclairant, des démonstrations présentes dans ce chapitre.

La première notion à introduire est la notion d'isomorphisme entre formes bilinéaires.

6.1. Définition. Soit k un corps. Deux formes bilinéaires $b_1 : V_1 \times V_1 \longrightarrow k$ et $b_2 : V_2 \times V_2 \longrightarrow k$ sont dites *isomorphes* s'il existe un isomorphisme d'espaces vectoriels $u : V_1 \xrightarrow{\sim} V_2$ tel que

$$b_2\big(u(x), u(y)\big) = b_1(x, y), \qquad \text{pour tous } x, y \in V_1.$$

On le note $(V_1, b_1) \simeq (V_2, b_2)$, voire $b_1 \simeq b_2$ s'il n 'y a pas de confusion possible.

En utilisant la définition, on voit facilement que, pour toutes formes bilinéaires $b : V \times V \longrightarrow k$, $b_i : V_i \times V_i \longrightarrow k$, $i \in [\![1, 3]\!]$,

(1) on a $b \simeq b$;

(2) si $b_1 \simeq b_2$, alors $b_2 \simeq b_1$;

(3) si $b_1 \simeq b_2$ et $b_2 \simeq b_3$, alors $b_1 \simeq b_3$.

6.2. Remarque. Si V_1 et V_2 sont de dimension finie, et si $(V_1, b_1) \simeq (V_2, b_2)$, alors V_1 et V_2 ont nécessairement même dimension.

En dimension finie, on a une caractérisation concrète en termes matriciels. Pour l'énoncer, introduisons une autre définition.

6.3. Définition. Soit k un corps. Deux matrices $M_1, M_2 \in \mathrm{M}_n(k)$ sont *congruentes* s'il existe $P \in \mathrm{GL}_n(k)$ telle que $P^t M_1 P = M_2$

On peut alors démontrer le résultat suivant.

6.4. Lemme. *Soit k un corps arbitraire, soient V_1, V_2 deux k-espaces vectoriels de dimension finie, et soient deux formes bilinéaires $b_1 : V_1 \times V_1 \longrightarrow k$ et $b_2 : V_2 \times V_2 \longrightarrow k$. Alors, les propriétés suivantes sont équivalentes :*

(1) *on a $(V_1, b_1) \simeq (V_2, b_2)$;*

(2) *pour toute base $\mathscr{B}_1$ de V_1 et toute base $\mathscr{B}_2$ de V_2, les matrices représentatives $\mathrm{Mat}(b_1 ; \mathscr{B}_1)$ et $\mathrm{Mat}(b_2 ; \mathscr{B}_2)$ sont congruentes ;*

(3) *il existe une base $\mathscr{B}_1$ de V_1 et une base $\mathscr{B}_2$ de V_2 telles que les matrices représentatives $\mathrm{Mat}(b_1 ; \mathscr{B}_1)$ et $\mathrm{Mat}(b_2 ; \mathscr{B}_2)$ sont congruentes.*

Notation. Si $a_1, \ldots, a_n \in k$, on note $\langle a_1, \ldots, a_n \rangle$ l'unique forme bilinéaire sur k^n dont la matrice représentative dans la base canonique est la matrice

$$\begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix}.$$

Autrement dit, c'est la forme bilinéaire symétrique

$$\begin{aligned} k^n \times k^n &\longrightarrow k \\ (x, y) &\longmapsto \sum_{i=1}^{n} a_i x_i y_i. \end{aligned}$$

Si $b : V \times V \longrightarrow k$ est une forme bilinéaire symétrique sur un k-espace vectoriel V de dimension finie, il est bien connu qu'il existe une base de V qui soit b-orthogonale (lorsque $\mathrm{car}(k) \neq 2$). En termes d'isomorphisme, cela se retraduit comme suit : il existe $a_1, \ldots, a_n \in k$ tel que $b \simeq \langle a_1, \ldots, a_n \rangle$.

De même, on constate que V possède une base b-orthonormée si, et seulement si, on a $b \simeq \langle 1, \ldots, 1 \rangle$.

6.5. Remarque. Soient $a_1, \ldots, a_n \in k$. Pour toute permutation σ de l'ensemble $[\![1, n]\!]$, on a

$$\langle a_1, \ldots, a_n \rangle \simeq \langle a_{\sigma(1)}, \ldots, a_{\sigma(n)} \rangle.$$

En effet, il suffit de faire la permutation appropriée sur les vecteurs de la base canonique pour s'en convaincre.

On continue en introduisant la notion de somme orthogonale de formes bilinéaires.

6.6. Définition. Soient $b_i : V_i \times V_i \longrightarrow k, i \in [\![1, r]\!]$, r formes bilinéaires. La *somme orthogonale* de $b_1, \ldots, b_r$, notée $(V_1, b_1) \perp \cdots \perp (V_r, b_r)$, ou plus

sobrement $b_1 \perp \cdots \perp b_r$, est la forme bilinéaire sur $V_1 \times \cdots \times V_r$ définie par

$$\begin{aligned}(V_1 \times \cdots \times V_r) \times (V_1 \times \cdots \times V_r) &\longrightarrow k \\ \big((x_1, \ldots, x_r), (y_1, \ldots, y_r)\big) &\longmapsto \sum_{i=1}^{r} b_i(x_i, y_i).\end{aligned}$$

La proposition suivante résume les principales propriétés de la somme orthogonale.

6.7. Proposition. *Soient $b : V \times V \longrightarrow k$, $b_i : V_i \times V_i \longrightarrow k$, $i \in [\![1, r]\!]$, des formes bilinéaires. Alors :*

(1) *pour tout $j \in [\![1, r]\!]$, l'isomorphisme évident*

$$(V_1 \times \cdots \times V_j) \times (V_{j+1} \times \cdots \times V_r) \simeq V_1 \times \cdots \times V_r$$

induit un isomorphisme

$$(b_1 \perp \cdots \perp b_j) \perp (b_{j+1} \perp \cdots \perp b_r) \simeq b_1 \perp \cdots \perp b_r \, ;$$

(2) *l'isomorphisme évident $V_1 \times V_2 \simeq V_2 \times V_1$ induit un isomorphisme*

$$b_1 \perp b_2 \simeq b_2 \perp b_1 \, ;$$

(3) *si $b_1 \simeq b_2$, alors $b \perp b_1 \simeq b \perp b_2$;*

(4) *on suppose que V_1 et V_2 sont de dimension finie. Si $\mathscr{B}_1 = (e_1, \ldots, e_n)$ et $\mathscr{B}_2 = (e'_1, \ldots, e'_m)$ sont des bases de V_1 et V_2 respectivement, alors la matrice représentative de $b_1 \perp b_2$ dans la base*

$$\mathscr{B} = \big((e_1, 0), \ldots, (e_n, 0), (0, e'_1), \ldots, (0, e'_m)\big)$$

est

$$\begin{pmatrix} \mathrm{Mat}(b_1 \, ; \mathscr{B}_1) & 0 \\ 0 & \mathrm{Mat}(b_2 \, ; \mathscr{B}_2) \end{pmatrix}.$$

En particulier, pour tous $a_1, \ldots, a_n, a'_1, \ldots, a'_m \in k$, on a

$$\langle a_1, \ldots, a_n \rangle \perp \langle a'_1, \ldots, a'_m \rangle \simeq \langle a_1, \ldots, a_n, a'_1, \ldots, a'_m \rangle.$$

6.8. Remarque. Le dernier point se généralise bien entendu à r formes bilinéaires.

Notation. Si $r \geqslant 1$ est un entier, la somme orthogonale $b \perp \cdots \perp b$ est notée simplement $r \times b$.

On constate alors que la forme bilinéaire symétrique θ_b du théorème 4.5 n'est rien d'autre que $4 \times b$. On peut alors récrire la démonstration de ce théorème comme suit.

Le choix d'une base b-orthogonale $(e_1, \ldots, e_n)$ de V donne un isomorphisme

$$b \simeq \langle a_1, \ldots, a_n \rangle.$$

Comme a_i n'est rien d'autre que $b(e_i, e_i)$, on a $a_i > 0$ pour tout $i \in [\![1, n]\!]$.

Les propriétés de la somme orthogonale et de la relation d'isomorphisme entraînent alors la succession d'isomorphismes suivante :

$$\begin{aligned} 4\times b &\simeq \langle a_1,\dots,a_n\rangle \perp \cdots \perp \langle a_1,\dots,a_n\rangle \\ &\simeq \langle a_1,\dots,a_n,\dots,a_1,\dots,a_n\rangle \\ &\simeq \langle a_1,a_1,a_1,a_1,\dots,a_n,a_n,a_n,a_n\rangle \\ &\simeq \langle a_1,a_1,a_1,a_1\rangle \perp \cdots \perp \langle a_n,a_n,a_n,a_n\rangle. \end{aligned}$$

Or, les calculs faits en fin de démonstration du théorème 4.5 montrent que l'on a un isomorphisme $\langle r,r,r,r\rangle \simeq \langle 1,1,1,1\rangle$ pour tout rationnel $r>0$.

On a ainsi

$$4\times b \simeq \langle 1,1,1,1\rangle \perp \cdots \perp \langle 1,1,1,1\rangle \simeq \langle 1,\dots,1\rangle,$$

ce qui revient à dire que V^4 possède une base orthonormée pour la forme bilinéaire $4\times b$.

Il nous reste à comprendre d'où peut bien sortir l'isomorphisme susmentioné. Pour cela, nous allons introduire le concept de forme bilinéaire symétrique multiplicative.

Notation. Si $b : V\times V \longrightarrow k$ est une forme bilinéaire symétrique, pour tout $\lambda\in k^\times$, on note λb la forme bilinéaire symétrique

$$\begin{aligned} \lambda b : V\times V &\longrightarrow k \\ (x,y) &\longmapsto \lambda b(x,y). \end{aligned}$$

Rappelons qu'une forme bilinéaire $b : V\times V\longrightarrow k$ est dite *anisotrope* si l'on a $b(x,x)\neq 0$ pour tout $x\in V\setminus\{0\}$, et *isotrope* sinon.

On a alors la définition suivante.

6.9. Définition. Soit k un corps dans lequel $2\neq 0$, et soit $b : V\times V\longrightarrow k$ une forme bilinéaire symétrique non dégénérée sur un k-espace vectoriel de dimension finie V.

Si b est isotrope, on dit que b est *multiplicative* si $\dim(V)$ est paire, et il existe un sous-espace [4] W de V tel que $\dim_k(W)=\dfrac{\dim_k(V)}{2}$ et $b(x,y)=0$, pour tous $x,y\in W$.

Si b est anisotrope, on dit que b est *multiplicative* si pour tout $\lambda\in k^\times$, on a

$$b\simeq \lambda b \iff \text{il existe } x_0\in V \text{ tel que } b(x_0,x_0)=\lambda.$$

6.10. Remarque. L'expression « forme multiplicative » est justifiée par le constat suivant.

4. Une telle forme bilinéaire est dite *hyperbolique*.

Supposons pour simplifier que b est anisotrope. Soit $x \in V \setminus \{0\}$, et choisissons un isomorphisme $u_x : V \xrightarrow{\sim} V$ tel que $b(x,x)b(y,z) = b\big(u_x(y), u_x(z)\big)$ pour tous $y, z \in V$ (un tel isomorphisme existe puisque b est multiplicative). En particulier, on a

$$b(x,x)b(y,y) = b\big(u_x(y), u_x(y)\big), \quad \text{pour tout } y \in V,$$

et l'ensemble $D(b) = \{b(x,x) \mid x \in V \setminus \{0\}\}$ est stable par produit.

Notons que si b est isotrope, on a $b \simeq m \times \langle 1, -1\rangle$, où $m = \dfrac{\dim_k(V)}{2}$ d'après le théorème 3.4. Or, on peut montrer que pour tout $\lambda \in k^\times$, on a un isomorphisme $\langle 1, -1\rangle \simeq \lambda\langle 1, -1\rangle$. On en déduit alors que $b \simeq \lambda b$ pour tout $\lambda \in k^\times$ dans ce cas.

6.11. Exemple. La forme $(k, \langle 1\rangle)$ est multiplicative. Elle est clairement anisotrope, et pour tout $x \in k^\times$, on a

$$x^2\langle 1\rangle = \langle x^2\rangle \simeq \langle 1\rangle,$$

le dernier isomorphisme étant induit par la multiplication par x dans k. Réciproquement, supposons que $\lambda\langle 1\rangle = \langle\lambda\rangle \simeq \langle 1\rangle$. Les automorphismes de l'espace vectoriel k étant les homothéties $x\mathrm{Id}, x \in k^\times$, en utilisant la définition d'isomorphisme de formes bilinéaires, on constate alors que λ est un carré dans k.

6.12. Théorème. *Soit $b : V \times V \longrightarrow k$ une forme multiplicative. Alors, pour tout $d \in k^\times$, la forme bilinéaire $b \perp db$ est multiplicative.*

6.13. Remarque. Le lecteur pourra retrouver toutes les notions et résultats introduits ici dans [26], par exemple.

L'exemple et le théorème précédents appliqué à $b = \langle 1\rangle$ et à $d = 1$, montrent que $\langle 1, 1\rangle$ est multiplicative. Une utilisation répétée de ce même théorème montre alors plus généralement que $b_m = 2^m \times \langle 1\rangle$ est multiplicative.

Comme les valeurs $b_m(x,x)$ sont exactement les sommes de 2^m carrés de k, on en déduit que si a est somme de 2^m carrés, alors

$$2^m \times \langle a\rangle = a(2^m \times \langle 1\rangle) \simeq 2^m \times \langle 1\rangle.$$

L'isomorphisme utilisé dans la démonstration du théorème 4.5 s'explique alors par le théorème des quatre carrés de Lagrange et l'isomorphisme précédent pour $m = 2$. La multiplicativité de b_m implique également que le produit de deux sommes de 2^m carrés est une somme de 2^m carrés, ce qui explique l'existence de l'identité d'Euler.

Nous continuons en réinterprétant le calcul d'une matrice de $\mathscr{S}_3(\mathbb{Q})$ de polynôme minimal $X(X^2-3)$.

Si $f = X^2 - 3$, et si $t : \mathbb{Q}[X]/(f) \longrightarrow \mathbb{Q}$ est la forme linéaire du lemme 4.6, rappelons que $b_t \simeq \langle 2, 6\rangle$. Comme on l'a déjà constaté, si V admet une base orthonormée pour une forme bilinéaire symétrique $b : V \times V \longrightarrow \mathbb{Q}$, alors le déterminant de la matrice représentative de b dans n'importe quelle base de V est un carré non nul.

En particulier, si $\langle a_1, \ldots, a_n\rangle \simeq n\times\langle 1\rangle$, alors $a_1 \cdots a_n$ est un carré non nul.

Or, si $t_i : \mathbb{Q}[X]/(f_i) \longrightarrow \mathbb{Q}$ est une forme linéaire pour tout $i \in [\![1, r]\!]$, et si $(t_1 \oplus \cdots \oplus t_r) : \mathbb{Q}[X]/(f_1) \times \cdots \times \mathbb{Q}[X]/(f_r) \longrightarrow \mathbb{Q}$ est la forme linéaire définie par

$$(t_1 \oplus \cdots \oplus t_r)(x_1, \ldots, x_r) = \sum_{i=1}^{r} t_i(x_i), \quad \text{pour tout } x_i \in \mathbb{Q}[X]/(f_i),$$

il est facile de voir que $b_{t_1\oplus\cdots\oplus t_r} = b_{t_1} \perp \cdots \perp b_{t_r}$.

Cela explique le choix de la forme linéaire $s : \mathbb{Q} \times \mathbb{Q}[X]/(f) \longrightarrow \mathbb{Q}$. On a en fait pris $s = (3\mathrm{Id}_{\mathbb{Q}}) \oplus t$, afin d'obtenir

$$b_s = b_{3\mathrm{Id}_{\mathbb{Q}}\oplus t} = \langle 3\rangle \perp b_t \simeq \langle 3\rangle \perp \langle 2, 6\rangle \simeq \langle 3, 2, 6\rangle.$$

Remarquons que l'on a bien $3\times2\times6 = 6^2$. Notons que le théorème 6.12 implique que pour tout $d \in \mathbb{Q}^\times$, et pour tout $r \in \mathbb{Q}^\times$ de la forme $r = a^2 + db^2$, on a

$$\langle r, rd\rangle = r\langle 1, d\rangle \simeq \langle 1, d\rangle,$$

ce qui explique les égalités matricielles un peu sorties du chapeau dans le paragraphe 4.

Mais alors, on a

$$\begin{aligned}
\langle 3, 2, 6\rangle &\simeq \langle 2, 3, 6\rangle \\
&\simeq \langle 2\rangle \perp (1^2 + 2\cdot1^2)\langle 1, 2\rangle \\
&\simeq \langle 2\rangle \perp \langle 1, 2\rangle \\
&\simeq \langle 2, 1, 2\rangle \\
&\simeq \langle 1, 2, 2\rangle \\
&\simeq \langle 1\rangle \perp (1^2 + 1^2)\langle 1, 1\rangle \\
&\simeq \langle 1\rangle \perp \langle 1, 1\rangle \\
&\simeq \langle 1, 1, 1\rangle.
\end{aligned}$$

Cela résume les calculs effectués à la fin du paragraphe 4 de manière plus conceptuelle. Notons néanmoins que ce point de vue ne donne pas une base orthonormée concrète. Pour en calculer une, il faut expliciter chaque isomorphisme, et c'est d'ailleurs exactement ce que l'on a fait.

À titre d'illustration des techniques introduites dans ce paragraphe, nous proposons la généralisation suivante du théorème 3.1.

6.14. Théorème. *Soit k un corps de caractéristique différente de 2. On suppose que -1 est somme de 2^r carrés de k.*

Alors, tout polynôme unitaire $f \in k[X]$ de degré $n \geqslant 1$ est le polynôme minimal d'une matrice de $\mathscr{S}_{2^r n}(k)$.

Démonstration. Écrivons $n = 2\ell$ ou $n = 2\ell - 1$, avec $\ell \geqslant 1$, selon la parité de n. Un examen attentif de la démonstration de la proposition 3.5 montre que la forme linéaire $s : k[X]/(f) \longrightarrow k$ définie dans l'énoncé de cette proposition vérifie

$$b_s \simeq m\times\langle 1\rangle \perp \ell\times\langle -1\rangle \quad \text{ou} \quad b_s \simeq \ell\times\langle 1\rangle \perp (\ell-1)\times\langle -1\rangle,$$

selon que n est pair ou impair.

Comme $2^r\times\langle 1\rangle$ est multiplicative, la remarque 6.10 montre que l'on a

$$2^r\times\langle 1\rangle \simeq (-1)(2^r\times\langle 1\rangle) = 2^r\times\langle -1\rangle,$$

que l'on soit dans le cas isotrope ou anisotrope.

Mais alors, $2^r\times b_s \simeq 2^r n\times\langle 1\rangle$. Une application de la proposition 2.3 avec la forme $s\oplus\cdots\oplus s : \big(k[X]/(f)\big)^{2^r n} \longrightarrow k$ donne alors le résultat souhaité. □

6.15. Remarque. On peut démontrer que si -1 est une somme de n carrés dans k, et si n est minimal, alors n est une puissance de 2 (cf. [26]).

7. Exercices

Le but de cette série d'exercices est d'utiliser les outils introduits dans ce chapitre pour donner une caractérisation des corps k pour lesquels toute matrice carrée est semblable à une matrice symétrique.

E1. Soit k un corps, et soit $b : V\times V \longrightarrow k$ une forme bilinéaire symétrique non dégénérée, où V est un k-espace vectoriel de dimension finie.

On rappelle que b est *alternée* si $b(x,x) = 0$, pour tout $x \in V$.

Dans cet exercice, on se propose de démontrer que si b est non alternée, alors V possède une base b-orthogonale (i.e. une base $\mathscr{B} = (e_1,\ldots,e_n)$ telle que $b(e_i,e_j) = 0$, dès que $i \neq j$).

a. Justifier l'existence de $e_1 \in V$ tel que $a = b(e_1,e_1) \neq 0$.

b. On pose $U = ke_1$. Justifier que $V = U \oplus U^\perp$, et que la restriction de la forme bilinéaire b à $U^\perp \times U^\perp$ est non dégénérée.

c. On suppose que $U^\perp \neq 0$, et que la restriction de b à $U^\perp \times U^\perp$ est alternée.

(i) Montrer que k est nécessairement de caractéristique 2.

(ii) Soit $e_2 \in U^\perp \setminus \{0\}$. Justifier l'existence d'un vecteur $e_3 \in U^\perp$ tel que $b(e_2, e_3) = a$.

(iii) Montrer que les vecteurs $e_1 + e_2 + e_3, e_1 + e_2, e_1 + e_3$ sont deux à deux orthogonaux.

d. À l'aide des questions précédentes, montrer soigneusement le résultat souhaité par récurrence.

E2. Soit k un corps. Montrer l'équivalence des propriétés suivantes :

(i) pour tout $n \geqslant 1$, toute matrice de $\mathrm{M}_n(k)$ est semblable à une matrice symétrique ;

(ii) toute matrice compagnon à coefficients dans k est semblable à une matrice symétrique ;

(iii) tout polynôme $f \in k[X]$ unitaire de degré $d \geqslant 1$ est le polynôme minimal d'une matrice de $\mathscr{S}_d(k)$.

E3. Soit k un corps de caractéristique différente de 2. En s'appuyant sur les résultats de ce chapitre et les exercices précédents, montrer l'équivalence des propriétés suivantes :

(i) pour tout $n \geqslant 1$, toute matrice de $\mathrm{M}_n(k)$ est semblable à une matrice symétrique ;

(ii) toute matrice de $\mathrm{M}_2(k)$ est semblable à une matrice symétrique ;

(iii) on a $-1 \in k^{\times 2}$.

E4. Soit k un corps arbitraire, et soient E_1, E_2 deux k-algèbres commutatives, associatives et unitaires, de dimension finie sur k. Pour $i = 1, 2$, soit $s_i : E_i \longrightarrow k$ une forme linéaire, et soit

$$\begin{aligned} s_1 \oplus s_2 \, : \, E_1 \times E_2 &\longrightarrow \quad k \\ (x_1, x_2) &\longmapsto s_1(x_1) + s_2(x_2). \end{aligned}$$

a. Montrer que $b_{s_1 \oplus s_2}$ est non dégénérée si, et seulement si, b_{s_1} et b_{s_2} le sont.

b. Soit $m \geqslant 1$, et soit $s_1 : k[X]/(X^m) \longrightarrow k$ l'unique forme linéaire telle que

$$s_1(\overline{X^i}) = 1, \quad \text{pour tout } i \in [\![0, m-1]\!].$$

Montrer que b_{s_1} est non dégénérée.

c. Soit $g = X^d + a_{d-1}X^{d-1} + \cdots + a_1 X + a_0 \in k[X]$, avec $d \geqslant 1$. On suppose que $a_0 \in k^\times$. Soit $s_2 : k[X]/(g) \longrightarrow k$ l'unique forme linéaire telle que

$$s_2(\overline{1}) = 1, s_2(\overline{X}) = \cdots = s_2(\overline{X^{d-1}}) = 0.$$

Montrer que b_{s_2} est non dégénérée.

d. En déduire soigneusement que, pour tout polynôme unitaire $f \in k[X]$, il existe une forme linéaire $s : k[X]/(f) \longrightarrow k$ telle que b_s soit non dégénérée et non alternée.

E5. Soit k un corps de caractéristique 2. En s'appuyant sur les exercices précédents, montrer l'équivalence des propriétés suivantes :

(i) pour tout $n \geqslant 1$, toute matrice de $\mathrm{M}_n(k)$ est semblable à une matrice symétrique ;

(ii) toute matrice de $\mathrm{M}_2(k)$ est semblable à une matrice symétrique ;

(iii) tout élément de k est un carré dans k (c'est-à-dire que k est un corps parfait).

Sixième partie

Représentations : agissons dans l'espace !

Chapitre XXXIV

Représentations linéaires des groupes finis : introduction

Il est toujours éclairant de réinterpréter les définitions de la théorie des représentations dans le contexte de la théorie des modules, même si ce n'est pas nécessaire. Nous commençons donc par quelques compléments sur les modules, qui nous serviront par la suite.

1. Modules simples, modules indécomposables

Dans cette section, A désigne un anneau non nécessairement commutatif.

On s'intéresse tout d'abord à la notion de sous-modules en somme directe.

1.1. Définition. Soit M un A-module et soit $(M_i)_{i\in I}$ une famille de *sous-modules* de M. On note $\sum_{i\in I} M_i$ le sous-module de M engendré par $\bigcup_{i\in I} M_i$.

On dit que les sous-modules M_i sont en *somme directe (interne)* si la propriété suivante est satisfaite : pour tous $x_i \in M_i$, $i \in I$ presque tous nuls,

$$\sum_{i\in I} x_i = 0 \quad \Longrightarrow \quad x_i = 0, \text{ pour tout } i \in I.$$

On dit que M est la *somme directe (interne) des sous-modules* M_i, $i \in I$ si les deux conditions suivantes sont vérifiées :

(1) on a $M = \sum_{i\in I} M_i$;

(2) les modules M_i, $i \in I$ sont en somme directe.

De manière équivalente, M est la somme directe des sous-modules M_i, $i \in I$ si pour tout $x \in M$, il existe une *unique décomposition*

$$x = \sum_{i \in I} x_i, \quad x_i \in M_i \text{ presque tous nuls.}$$

Dans ce cas, on le note $M = \bigoplus_{i \in I} M_i$.

1.2. Remarque. Si A est un corps K, on retrouve la notion de somme directe interne de sous-espaces vectoriels.

Nous continuons par un lemme qui résume les principales propriétés de la somme directe interne.

La propriété (4) qui y figure est particulièrement importante, et sera systématiquement utilisée dans les chapitres suivants sans référence particulière à ce lemme.

1.3. Lemme. *Soit M un A-module, et soit $(M_i)_{i \in I}$ une famille de sous-modules de M. Alors, on a les propriétés suivantes :*

(1) *les sous-modules M_i, $i \in I$ sont en somme directe si, et seulement si,*

$$\sum_{i \in I} M_i = \bigoplus_{i \in I} M_i ;$$

(2) *on a $M = \bigoplus_{i \in I} M_i$ si, et seulement si, l'application*

$$f : \coprod_{i \in I} M_i \longrightarrow M$$
$$(x_i)_{i \in I} \longmapsto \sum_{i \in I} x_i$$

est un isomorphisme de A-modules ;

(3) *soit $(M_i)_{i \in I}$ une famille de A-modules, et soit*

$$M_i' = \{(0, \ldots, 0, x_i, 0, \ldots) \mid x_i \in M_i\}.$$

Alors, on a

$$\coprod_{i \in I} M_i = \bigoplus_{i \in I} M_i' ;$$

(4) *si $M = \bigoplus_{i \in I} M_i$, et si $f_i : M_i \longrightarrow M'$ est une application linéaire, alors il existe une unique application linéaire $f : M \longrightarrow M'$ vérifiant*

$$f_{|M_i} = f_i, \quad \text{pour tout } i \in I.$$

Démonstration. Les points (1), (2) et (3) proviennent de la définition. Gardons les notations de l'énoncé, et démontrons (4). Supposons qu'une telle

application $f : M \longrightarrow M'$ existe. Par hypothèse, tout élément de M peut s'écrire de manière unique sous la forme

$$\sum_{i\in I} x_i, \quad x_i \in M_i \text{ presque tous nuls.}$$

Puisque f est linéaire, on a

$$f\Big(\sum_{i\in I} x_i\Big) = \sum_{i\in I} f(x_i) = \sum_{i\in I} f_i(x_i).$$

Il s'ensuit que si f existe, elle est définie par la formule

$$f\Big(\sum_{i\in I} x_i\Big) = \sum_{i\in I} f_i(x_i),$$

d'où l'unicité. Pour conclure, il suffit de montrer que l'application f définie de la manière ci-dessus est effectivement linéaire. Or, nous avons

$$\begin{aligned} f\Big(\sum_{i\in I} x_i + \sum_{i\in I} x'_i\Big) &= f\Big(\sum_{i\in I}(x_i + x'_i)\Big) \\ &= \sum_{i\in I} f_i(x_i + x'_i) \\ &= \sum_{i\in I} \big(f_i(x_i) + f_i(x'_i)\big) \\ &= \sum_{i\in I} f_i(x_i) + \sum_{i\in I} f_i(x'_i) \\ &= f\Big(\sum_{i\in I} x_i\Big) + f\Big(\sum_{i\in I} x'_i\Big). \end{aligned}$$

De même, pour tout $a \in A$, on a

$$\begin{aligned} f\Big(a\cdot\Big(\sum_{i\in I} x_i\Big)\Big) &= f\Big(\sum_{i\in I} a\cdot x_i\Big) \\ &= \sum_{i\in I} f(a\cdot x_i) \\ &= \sum_{i\in I} a\cdot f(x_i) \\ &= a\cdot \sum_{i\in I} f(x_i) \\ &= a\cdot f\Big(\sum_{i\in I} x_i\Big). \end{aligned}$$

Ainsi, f est bien linéaire, ce qui achève la démonstration. □

1.4. Définition. On dit qu'un A-module M est *indécomposable* s'il est non nul et si pour tous sous-modules P et Q de M, on a

$$M = P \oplus Q \quad \Longrightarrow \quad P = 0 \text{ ou } Q = 0.$$

On dit qu'un A-module M est *simple* s'il est non nul et si les seuls sous-modules de M sont 0 et M.

1.5. Exemples

(1) Un A-module simple est bien évidemment indécomposable.

(2) Soit A un anneau intègre qui n'est pas un corps. Alors, le A-module A est indécomposable, mais n'est pas simple.

En effet, comme A n'est pas un corps, il existe un élément $a \in A \setminus \{0\}$ non inversible. L'idéal (a) est alors un sous-module de A distinct de $\{0\}$ et de A.

Supposons maintenant que $A = P \oplus Q$, où P et Q sont deux sous-modules de A, c'est-à-dire deux idéaux de A. Supposons que les sous-modules P et Q soient non nuls, et soient $x \in P \setminus \{0\}$ et $y \in Q \setminus \{0\}$. Comme P et Q sont des idéaux, on a $xy \in P \cap Q$. L'hypothèse entraîne alors que $xy = 0$, d'où une contradiction par intégrité de A, puisque x et y sont non nuls.

On finit cette section par un lemme dû à Schur.

1.6. Lemme (Schur). *Soient M et N deux A-modules simples. Alors, toute application A-linéaire $f : M \longrightarrow N$ est soit nulle, soit un isomorphisme. En particulier :*

(1) *si M et N sont simples et non isomorphes, alors* $\mathrm{Hom}_A(M,N) = 0$;

(2) *si M est un A-module simple,* $\mathrm{End}_A(M)$ *est un anneau à division ;*

(3) *on suppose que A est une K-algèbre, où K est un corps algébriquement clos, et que M un A-module simple de dimension finie sur K. Alors, on a* $\mathrm{End}_A(M) = K\mathrm{Id}_M$.

Démonstration. Soit $f : M \longrightarrow N$ une application A-linéaire. Supposons que f soit non nulle. Alors, $\mathrm{Ker}(f) \neq M$ et $\mathrm{Im}(f) \neq 0$. La simplicité de M et N montre alors respectivement que $\mathrm{Ker}(f) = 0$ et $\mathrm{Im}(f) = N$. Ainsi, f est un isomorphisme. Les points (1) et (2) en découlent immédiatement.

Montrons (3). Soit $f \in \mathrm{End}_A(M)$. Alors, f est en particulier un endomorphisme du K-espace vectoriel M, qui est de dimension finie. Puisque le corps K est algébriquement clos, f possède une valeur propre $\lambda \in K$. Mais alors, $f - \lambda\mathrm{Id}_M$ est un élément de $\mathrm{End}_A(M)$ non injectif. En particulier, ce n'est pas un isomorphisme. Comme M est simple, on obtient $f - \lambda\mathrm{Id}_M = 0$, d'où le résultat. □

2. Représentations linéaires : définitions et premiers exemples

Il est incontestable qu'une bonne façon[1] d'étudier un groupe est de le faire agir sur un ensemble, et nous espérons qu'à ce stade, le lecteur de cet ouvrage en soit déjà convaincu. Pour preuve, il suffit de recenser les résultats obtenus grâce à la théorie des actions de groupe dans la deuxième partie de ce livre.

Rappelons au passage que faire agir un groupe G sur un ensemble E non vide revient à se donner un morphisme de groupes $G \longrightarrow \mathfrak{S}(E)$. Bien que les propriétés du groupe symétrique soient bien connues, l'étude de l'ensemble de ces morphismes reste ardue.

L'idée centrale de la théorie des représentations linéaires est donc de rigidifier un peu la situation en imposant des contraintes supplémentaires sur l'action, et de se restreindre à l'étude des actions dites *linéaires*. Ce n'est d'ailleurs pas une restriction très contraignante, puisque l'on verra dans l'exemple 2.8 que toute action de G sur un ensemble E donne lieu à une action linéaire sur un espace vectoriel bien choisi.

2.1. Définition. Soit G un groupe et soit K un corps. Une action *linéaire* de G sur un K-espace vectoriel V est une action $G \times V \longrightarrow V$ telle que, pour tous v, v_1, $v_2 \in V$, tout $g \in G$ et tout $\lambda \in K$, on a

$$g{\cdot}(\lambda v) = \lambda(g{\cdot}v) \quad \text{et} \quad g{\cdot}(v_1 + v_2) = g{\cdot}v_1 + g{\cdot}v_2.$$

On dit aussi que G agit *linéairement* sur V.

Notons que si G agit linéairement sur le K-espace vectoriel V, l'application

$$\begin{aligned} \sigma_g : V &\longrightarrow V \\ v &\longmapsto g{\cdot}v \end{aligned}$$

est par définition un endomorphisme de V pour tout $g \in G$. Comme cette application est bijective, cela montre que le morphisme de groupes associé à l'action de G sur V est en fait un morphisme $G \longrightarrow \mathrm{GL}(V)$. Inversement, si $\rho : G \longrightarrow \mathrm{GL}(V)$ est un morphisme de groupes, l'action associée, définie par

$$\begin{aligned} G \times V &\longrightarrow V \\ (g, v) &\longmapsto g{\cdot}v = \rho(g)(v) \end{aligned}$$

est une action linéaire.

1. Voire la meilleure...

Autrement dit, on a une correspondance bijective entre l'ensemble des actions linéaires de G sur un K-espace vectoriel V et l'ensemble des morphisme des groupes $G \longrightarrow \mathrm{GL}(V)$. L'avantage d'étudier les actions linéaires de G est de permettre l'utilisation de tout l'arsenal de l'algèbre linéaire, en particulier la réduction des endomorphismes (lorsque V est de dimension finie sur K).

Cela amène tout naturellement à la définition suivante.

2.2. Définition. Une *représentation linéaire* d'un groupe G sur un corps K est un K-espace vectoriel muni d'une action linéaire de G. De manière équivalente, une représentation linéaire de G sur K est un morphisme de groupes $\rho : G \longrightarrow \mathrm{GL}(V)$, où V est un K-espace vectoriel.

On dit qu'une représentation linéaire V de G est *fidèle* si G agit fidèlement sur V, i.e. si le morphisme $\rho : G \longrightarrow \mathrm{GL}(V)$ correspondant est injectif.

Si V est de dimension finie, l'entier $\dim_K(V)$ est appelé le *degré* de la représentation.

Lorsque V est une droite, i.e. lorsque $\dim_K(V) = 1$, on a un isomorphisme de groupes canonique

$$\begin{aligned} K^\times &\xrightarrow{\sim} \mathrm{GL}(V) \\ \lambda &\longmapsto \lambda \mathrm{Id}_V, \end{aligned}$$

et l'on identifie un morphisme de groupes $G \longrightarrow \mathrm{GL}(V)$ à un morphisme de groupes $G \longrightarrow K^\times$.

Un morphisme $G \longrightarrow K^\times$ (i.e une représentation linéaire de degré 1) est appelé un *caractère linéaire* de G (à valeurs dans K). Cette définition est en accord avec la définition XII-3.1 (où les caractères linéaires considérés sont à valeurs dans $\mathbb{C}$).

Le K-espace vectoriel K, muni de l'action triviale de G, est appelé la *représentation triviale* de G sur K. Elle correspond au morphisme trivial

$$\mathbf{1} : G \longrightarrow K^\times.$$

2.3. Remarque. La construction de représentations d'un groupe G peut être grandement facilitée si l'on connaît une présentation de G par générateurs et relations. En effet, si $G = \langle X \mid R \rangle$, tout morphisme $G \longrightarrow \mathrm{GL}(V)$ est défini de manière unique par le choix d'une famille $(f_x)_{x \in X}$ d'éléments de $\mathrm{GL}(V)$ vérifiant les relations R (cf. proposition IX-4.24).

Dans la suite, G désignera un groupe et K désignera un corps. On abrégera souvent l'expression « représentation linéaire de G sur K » par « représentation de G ».

De plus, on utilisera indifféremment les deux points de vue équivalents, et l'on pourra dire « soit V une représentation de G », aussi bien que « soit $\rho : G \longrightarrow \mathrm{GL}(V)$ une représentation de G », ou encore parfois « soit (V, ρ) une représentation de G », selon le contexte.

2.4. Remarque. Si $V = K^n$, on a un isomorphisme canonique

$$\mathrm{GL}(K^n) \simeq \mathrm{GL}_n(K).$$

Dans ce cas, on identifiera souvent une représentation $G \longrightarrow \mathrm{GL}(K^n)$ et le morphisme associé $G \longrightarrow \mathrm{GL}_n(K)$.

Un morphisme $G \longrightarrow \mathrm{GL}_n(K)$ sera appelé une *représentation matricielle* de G.

Si V est une représentation de G, il est facile de voir que le sous-ensemble V^G des vecteurs de V qui sont fixés par G est un sous-espace vectoriel de V. Le lemme suivant permet de calculer sa dimension, modulo certaines hypothèses sur G et K, qui reviendront constamment dans ce chapitre.

2.5. Lemme. *On suppose que G est fini, et que $\mathrm{car}(K) \nmid |G|$. Soit V une représentation de G de degré fini. Alors, l'application linéaire*

$$\begin{aligned} p_G : V &\longrightarrow V \\ v &\longmapsto \frac{1}{|G|} \sum_{g \in G} g \cdot v \end{aligned}$$

est un projecteur sur V^G. En particulier, on a

$$\mathrm{tr}(p_G) = \dim_K(V^G) \cdot 1_K.$$

Démonstration. Commençons par remarquer que pour tout vecteur $v \in V$ et tout $g \in G$, on a

$$g \cdot p_G(v) = g \cdot \Big(\frac{1}{|G|} \sum_{h \in G} h \cdot v \Big) = \frac{1}{|G|} \sum_{h \in G} gh \cdot v = p_G(v),$$

la dernière égalité étant obtenue grâce au fait que la translation à gauche par g dans G est bijective. On a donc $p_G(V) \subset V^G$. Par ailleurs, il est clair que $v = p_G(v) \in p_G(V)$ pour tout $v \in V^G$, et par conséquent

$$p_G(V) = V^G.$$

Montrons que p_G est un projecteur. On vient de voir que $p_G(v) \in V^G$ et que la restriction de p_G à V^G est l'identité. On a donc $p_G\big(p_G(v)\big) = p_G(v)$, pour tout $v \in V$, d'où le résultat.

Le dernier point vient du fait bien connu que, si $p \in \mathscr{L}(V)$ est un projecteur, on a

$$\mathrm{tr}(p) = \mathrm{rg}(p) \cdot 1_K.$$

Pour le voir, il suffit de se rappeler que si $p \in \mathscr{L}(V)$ est un projecteur, alors on a

$$V = \mathrm{Im}(p) \oplus \mathrm{Ker}(p)$$

par le corollaire V-1.5. Par conséquent, dans une base $\mathbf{e}$ adaptée à cette décomposition en somme directe, la matrice représentative de p est

$$\mathrm{Mat}(p\,;\mathbf{e}) = \begin{pmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{pmatrix},$$

où $r = \dim_K\big(\mathrm{Im}(p)\big)$. Cela achève la démonstration. □

2.6. Remarque. Si $\mathrm{car}(K) = 0$, la dernière partie du lemme montre que l'on a $\mathrm{tr}(p_G) = \dim_K(V^G)$.

Nous allons maintenant donner une première famille d'exemples, qui montre comment la théorie des actions de groupes peut être englobée dans celle des représentations linéaires.

Si E est un ensemble non vide, on note $K[E]$ l'ensemble des fonctions de E dans K à support fini. C'est un sous-espace vectoriel du K-espace vectoriel $\mathscr{F}(E, K)$ des fonctions de E dans K.

Pour tout $x \in E$, on note $e_x \in K[E]$ la fonction définie par

$$e_x(x') = \begin{cases} 1 & \text{si } x' = x \\ 0 & \text{sinon.} \end{cases}$$

2.7. Lemme. *La famille $(e_x)_{x\in E}$ est une K-base de $K[E]$. De plus, pour tout $f \in K[E]$, on a*

$$f = \sum_{x\in E} f(x) e_x.$$

Enfin, si E est fini, on a

$$\dim_K(K[E]) = |E|.$$

Démonstration. Soit $(\lambda_x)_{x\in E}$ une famille d'éléments de K presque tous nuls. Pour tout $x' \in E$, on a

$$\Big(\sum_{x\in E} \lambda_x e_x\Big)(x') = \sum_{x\in E} \lambda_x e_x(x') = \lambda_{x'}.$$

En particulier, si $\sum_{x\in E} \lambda_x e_x = 0$, on a $\lambda_x = 0$ pour tout vecteur $x \in E$, et la famille $(e_x)_{x\in E}$ est libre.

Soit maintenant $f \in K[E]$. Par hypothèse, la famille $\big(f(x)\big)_{x\in E}$ est à support fini, et le calcul fait au début montre que l'on a

$$\Big(\sum_{x\in E} f(x)e_x\Big)(x') = f(x'), \quad \text{pour tout } x' \in E.$$

Autrement dit, $f = \sum_{x\in E} f(x)e_x$, et la famille $(e_x)_{x\in E}$ est génératrice. Par conséquent, la famille $(e_x)_{x\in E}$ est une K-base de $K[E]$. □

2.8. Exemple. On suppose que le groupe G agit sur E. Pour tout $g \in G$ et tout $f \in K[E]$, on pose

$$(g{\cdot}f)(x) = f(g^{-1}{\cdot}x).$$

Clairement, $g{\cdot}f \in K[E]$, puisque f est à support fini. On vérifie aisément que l'application

$$\begin{aligned} G \times K[E] &\longrightarrow K[E] \\ (g, f) &\longmapsto g{\cdot}f \end{aligned}$$

est une action linéaire de G sur $K[E]$. Ainsi, $K[E]$ est une représentation linéaire de G.

Les définitions donnent immédiatement l'égalité

$$g{\cdot}e_x = e_{g\cdot x}, \quad \text{pour tout } g \in G \text{ et tout } x \in E.$$

Autrement dit, un élément $g \in G$ permute les éléments de la base $(e_x)_{x\in E}$ de la même manière que g permute les éléments de E.

L'action de G sur $K[E]$ peut aussi se décrire de la manière suivante :

$$g{\cdot}\Big(\sum_{x\in E} \lambda_x e_x\Big) = \sum_{x\in E} \lambda_x e_{g\cdot x} = \sum_{x\in E} \lambda_{g^{-1}\cdot x} e_x.$$

Remarquons que la représentation $K[E]$ est fidèle si, et seulement si, G agit fidèlement sur E.

En effet, soit $g \in G$. Alors, l'application

$$\begin{aligned} K[E] &\longrightarrow K[E] \\ f &\longmapsto g{\cdot}f \end{aligned}$$

est l'identité de $K[E]$ si, et seulement si, on a

$$g{\cdot}e_x = e_x, \quad \text{pour tout } x \in E.$$

Puisque $g{\cdot}e_x = e_{g\cdot x}$, cela équivaut à

$$g{\cdot}x = x, \quad \text{pour tout } x \in E.$$

On en déduit immédiatement le résultat voulu.

Ces considérations justifient la définition suivante.

2.9. Définition. Une représentation linéaire de G de la forme $K[E]$, où E est un ensemble non vide sur lequel G agit, s'appelle une *représentation de permutation*.

2.10. Exemple. Le groupe G agit sur lui-même par translation à gauche. On obtient alors une représentation linéaire $K[G]$, où l'action du groupe G sur $K[G]$ est donnée par

$$(g\cdot f)(h) = f(g^{-1}h), \quad \text{pour tous } g,\, h \in G \text{ et tout } f \in K[G].$$

2.11. Définition. La représentation de permutation associée à l'action de G sur lui-même par translation à gauche est appelée la *représentation régulière* de G. Lorsque G est fini, son degré est égal à $|G|$.

2.12. Remarque. Puisque l'action de G sur lui-même par translation à gauche est fidèle, la représentation régulière de G est une représentation fidèle, par l'exemple 2.8.

Continuons à étudier la représentation $K[E]$.

2.13. Exemple. On suppose que G agit sur E, et l'on considère la représentation $K[E]$.

Déterminons $K[E]^G$. Une fonction $f \in K[E]$ est fixe sous l'action de G si, et seulement si, $f(g^{-1}\cdot x) = f(x)$ pour tout $g \in E$ et tout $x \in E$, ou encore si, et seulement si,

$$f(g\cdot x) = f(x), \quad \text{pour tout } g \in G \text{ et tout } x \in E.$$

Autrement dit, les éléments de $K[E]^G$ sont les fonctions constantes sur chaque G-orbite de E.

Supposons maintenant que E soit fini, et notons Ω l'ensemble des G-orbites. C'est donc un ensemble fini. Pour tout $\omega \in \Omega$, posons

$$e_\omega = \sum_{x\in\omega} e_x,$$

et choisissons un élément $x_\omega \in \omega$. Clairement, $e_\omega \in K[E]^G$ pour tout $\omega \in \Omega$. De plus, la famille $(e_\omega)_{\omega\in\Omega}$ est libre, puisque $(e_x)_{x\in E}$ est une famille libre de $K[E]$. Enfin, un élément $f \in K[E]^G$ s'écrit

$$f = \sum_{\omega\in\Omega} f(x_\omega)e_\omega,$$

et $(e_\omega)_{\omega\in\Omega}$ est en fait une base de $K[E]^G$.

Lorsque E est fini, on a donc

$$\dim_K \left(K[E]^G\right) = |\Omega|.$$

Nous allons maintenant calculer cette dimension d'une autre manière, en utilisant le lemme 2.5. On suppose toujours que E est fini.

Soit $\rho : G \longrightarrow \mathrm{GL}(K[E])$ le morphisme associé à l'action de G sur $K[E]$. On commence par calculer $\mathrm{tr}\big(\rho(g)\big)$ pour tout $g \in G$. Par définition, on a

$$\rho(g)(e_x) = g{\cdot}e_x = e_{g\cdot x}, \quad \text{pour tout } g \in G \text{ et tout } x \in E.$$

On a alors

$$\mathrm{Mat}\big(\rho(g)\,;\,(e_x)_{x\in E}\big) = (\delta_{g^{-1}\cdot x,x'})_{x,x'\in E}.$$

Les éléments diagonaux sont donc les éléments $\delta_{g^{-1}\cdot x,x}$, $x \in E$. Or, on a

$$\delta_{g^{-1}\cdot x,x} = 1_K \iff x \in \mathrm{Fix}(g) = \{x \in E \mid g{\cdot}x = x\}.$$

Par conséquent, on a

$$\mathrm{tr}\big(\rho(g)\big) = |\mathrm{Fix}(g)|{\cdot}1_K, \quad \text{pour tout } g \in G.$$

Supposons enfin que G soit fini, et que $\mathrm{car}(K) \nmid |G|$. Le lemme 2.5 donne alors

$$\dim_K \left(K[E]^G\right) = \frac{1}{|G|}\sum_{g\in G} |\mathrm{Fix}(g)|.$$

En appliquant ce qui précède à $K = \mathbb{C}$, et en utilisant l'injectivité du morphisme canonique $\Theta_K : \mathbb{Z} \longrightarrow K$, on obtient le résultat suivant : pour tout groupe fini G agissant sur un ensemble fini non vide E, on a

$$|\Omega| = \frac{1}{|G|}\sum_{g\in G} |\mathrm{Fix}(g)|,$$

ce qui n'est rien d'autre que la formule de Burnside (proposition VII-1.15).

Nous allons maintenant établir un lien entre la théorie des représentations et la théorie des modules. Pour cela, nous allons tout d'abord définir une structure de K-algèbre associative unitaire sur $K[G]$, qui étend la loi de groupe de G.

Puisque $(e_g)_{g\in G}$ est une K-base de $K[G]$, il existe une unique application bilinéaire $K[G] \times K[G] \longrightarrow K[G]$ qui envoie le couple $(e_g, e_{g'})$ sur $e_{gg'}$ pour tous g, $g' \in G$. Le K-espace vectoriel $K[G]$, muni de cette forme bilinéaire, est donc une K-algèbre dont le produit est défini par

$$\Big(\sum_{g\in G}\lambda_g e_g\Big)\Big(\sum_{g\in G}\mu_g e_g\Big) = \sum_{g_1,g_2\in G}\lambda_{g_1}\mu_{g_2}e_{g_1g_2} = \sum_{g\in G}\Big(\sum_{h\in G}\lambda_h\mu_{h^{-1}g}\Big)e_g.$$

Notons que $K[G]$ est unitaire, d'élément unité e_{1_G}.

Enfin, la loi produit est associative. En effet, par bilinéarité, il suffit de vérifier l'égalité

$$(e_g e_h)e_k = e_g(e_h e_k), \quad \text{pour tous } g,\ h,\ k \in G,$$

qui découle de la définition du produit et de l'associativité de la loi de groupe de G.

Notons au passage que $K[G]$ étant une K-algèbre associative unitaire, c'est aussi un anneau.

2.14. Définition. Soit G un groupe et soit K un corps. La K-algèbre associative unitaire $K[G]$ est appelée *la K-algèbre du groupe G.*

Revenons maintenant au sujet de nos préoccupations. Si V est une représentation de G, on vérifie aisément que la loi externe

$$\begin{aligned} K[G] \times V &\longrightarrow V \\ (f, v) &\longmapsto f * v = \sum_{g \in G} f(g)(g \cdot v) \end{aligned}$$

induit sur le groupe abélien $(V, +)$ une structure de $K[G]$-module, que l'on notera $M(V)$. En d'autres termes, on a

$$\Big(\sum_{g \in G} \lambda_g e_g\Big) * v = \sum_{g \in G} \lambda_g (g \cdot v).$$

Inversement, si M est un $K[G]$-module, dont la loi externe est notée $*$, la loi externe

$$\begin{aligned} K \times M &\longrightarrow M \\ (\lambda, x) &\longmapsto (\lambda e_{1_G}) * x \end{aligned}$$

induit sur le groupe abélien $(M, +)$ une structure de K-espace vectoriel. On vérifie alors que l'action

$$\begin{aligned} G \times M &\longrightarrow M \\ (g, x) &\longmapsto g \bullet x = e_g * x \end{aligned}$$

est linéaire. On obtient ainsi une représentation de G, notée $V(M)$.

2.15. Lemme. *Les applications*

$$\begin{aligned} V &\longmapsto M(V) \\ V(M) &\longleftarrow M \end{aligned}$$

induisent une correspondance bijective entre l'ensemble des représentations de G et l'ensemble des $K[G]$-modules.

Démonstration. Soit V une représentation de G. Notons $M = M(V)$. Par définition, on a $(M,+) = (V,+)$. De plus, pour tout $\lambda \in K$, et tout $v \in M$, on a

$$(\lambda e_{1_G}) * v = \lambda(1_G \cdot v) = \lambda v,$$

et par conséquent, les K-espaces vectoriels $V(M)$ et V coïncident. Enfin, pour tout $g \in G$ et tout $v \in V(M)$, on a

$$g \bullet v = e_g * v = g \cdot v.$$

Bref, $V(M)$ est la représentation V.

Inversement, soit M un $K[G]$-module, dont la loi externe est notée $*$, et soit $V = V(M)$ la représentation correspondante. On a $(V,+) = (M,+)$, et la loi externe sur V est donc définie par

$$\lambda x = (\lambda e_{1_G}) * x,$$

pour tout $\lambda \in K$ et tout $x \in M$. De plus, pour tout $g \in G$ et tout $x \in M$, on a

$$g \bullet x = e_g * x.$$

Mais alors, on a

$$\lambda(g \bullet x) = (\lambda e_{1_G}) * (e_g * x) = \big((\lambda e_{1_G}) e_g\big) * x = (\lambda e_g) * x.$$

On en déduit alors aisément que la loi externe de $M(V)$ n'est rien d'autre que la loi externe de M. Ainsi, $M(V) = M$, d'où le résultat. □

2.16. Exemple. Soit G un groupe et soit K un corps. La loi produit induit sur l'anneau $K[G]$ une structure de $K[G]$-module, et la représentation de G correspondante n'est rien d'autre que la représentation régulière.

Nous allons maintenant définir la notion de morphisme de représentations. Si V_1 et V_2 sont deux représentations de G, on s'attend évidemment à ce qu'un morphisme de représentations $f : V_1 \longrightarrow V_2$ soit compatible avec les structures sous-jacentes, à savoir les structures d'espaces vectoriels et les actions de G.

On veut donc que $f : V_1 \longrightarrow V_2$ soit K-linéaire, et G-équivariante, i.e. vérifie

$$f(g \cdot v_1) = g \cdot f(v_1), \quad \text{pour tout } g \in G \text{ et tout } v_1 \in V_1.$$

Si V_i correspond à un morphisme $\rho_i : G \longrightarrow \mathrm{GL}(V_i)$, i.e. si l'action de G sur V_i est donnée par

$$g \cdot v_i = \rho_i(g)(v_i),$$

pour tout $g \in G$ et tout $v_i \in V_i$, cela revient à dire que $f : V_1 \longrightarrow V_2$ est une application linéaire vérifiant

$$f\big(\rho_1(g)(v_1)\big) = \rho_2(g)\big(f(v_1)\big), \quad \text{pour tout } g \in G \text{ et tout } v_1 \in V_1,$$

soit encore $$f \circ \rho_1(g) = \rho_2(g) \circ f, \quad \text{pour tout } g \in G.$$

Enfin, et c'est peut-être le point de vue le plus naturel, si l'on réinterprète une représentation V comme un $K[G]$-module, il est aisé de constater qu'une application K-linéaire vérifiant les conditions souhaitées n'est rien d'autre qu'une application $K[G]$-linéaire.

En résumé, on peut donc poser la définition suivante.

2.17. Définition. Un *morphisme de représentations* $f : V_1 \longrightarrow V_2$ est une application K-linéaire telle que

$$f(g \cdot v_1) = g \cdot f(v_1), \quad \text{pour tout } g \in G \text{ et tout } v_1 \in V_1.$$

Il revient au même de dire que l'application $f : V_1 \longrightarrow V_2$ est $K[G]$-linéaire.

De manière équivalente, un morphisme de représentations

$$f : (V_1, \rho_1) \longrightarrow (V_2, \rho_2)$$

est une application K-linéaire $f : V_1 \longrightarrow V_2$ vérifiant

$$f \circ \rho_1(g) = \rho_2(g) \circ f, \quad \text{pour tout } g \in G.$$

On note $\mathrm{Hom}_G(V_1, V_2)$ l'ensemble des morphismes de représentations linéaires $V_1 \longrightarrow V_2$. Si $V_1 = V_2 = V$, on le note $\mathrm{End}_G(V)$.

On définit de manière évidente la notion d'isomorphisme (resp. automorphisme) de représentations, et de représentations isomorphes (on dit aussi « équivalentes »). Notons que deux représentations sont isomorphes si, et seulement si, elles sont isomorphes comme $K[G]$-modules.

2.18. Remarque. Soit V un K-espace vectoriel de dimension finie $n \geqslant 1$. Alors, toute représentation $\rho : G \longrightarrow \mathrm{GL}(V)$ est isomorphe à une représentation $G \longrightarrow \mathrm{GL}(K^n)$, et donc à une représentation matricielle.

En effet, le choix d'une base de V fournit un isomorphisme

$$u : K^n \xrightarrow{\sim} V.$$

Définissons $\rho' : G \longrightarrow \mathrm{GL}(K^n)$ par

$$\rho'(g) = u^{-1} \circ \rho(g) \circ u, \quad \text{pour tout } g \in G.$$

Alors, ρ et ρ' sont isomorphes par construction.

2.19. Exemples

(1) Soit $n \geqslant 2$. Le morphisme de signature induit une représentation

$$\begin{aligned} \mathfrak{S}_n &\longrightarrow K^\times \\ \sigma &\longmapsto \varepsilon(\sigma) 1_K \end{aligned}$$

de degré 1, non triviale dès que $\mathrm{car}(K) \neq 2$.

(2) Soit E un ensemble non vide sur lequel G agit. Alors, l'action de G sur $K^{(E)}$ définie par

$$\begin{aligned} G \times K^{(E)} &\longrightarrow K^{(E)} \\ \big(g, (\lambda_x)_{x\in E}\big) &\longmapsto (\lambda_{g^{-1}\cdot x})_{x\in E} \end{aligned}$$

est linéaire. On obtient ainsi une représentation de G.

Les calculs faits dans l'exemple 2.8 montrent que l'application

$$\begin{aligned} K^{(E)} &\longrightarrow K[E] \\ (\lambda_x)_{x\in E} &\longmapsto \sum_{x\in E} \lambda_x e_x \end{aligned}$$

est un isomorphisme de représentations.

(3) Soit $n \geqslant 1$. Le groupe $\mathfrak{S}_n$ agit linéairement sur K^n par

$$\begin{aligned} \mathfrak{S}_n \times K^n &\longrightarrow K^n \\ \big(\sigma, (x_i)_{i\in[\![1,n]\!]}\big) &\longmapsto (x_{\sigma^{-1}(i)})_{i\in[\![1,n]\!]}. \end{aligned}$$

On obtient ainsi une représentation de permutation de degré n, qui n'est rien d'autre que la représentation de permutation associée à l'action naturelle de $\mathfrak{S}_n$ sur $[\![1,n]\!]$, à isomorphisme près.

Après identification de $\mathrm{GL}(K^n)$ avec $\mathrm{GL}_n(K)$, cette représentation correspond à la représentation matricielle

$$\begin{aligned} \rho : \mathfrak{S}_n &\longrightarrow \mathrm{GL}_n(K) \\ \sigma &\longmapsto M_\sigma, \end{aligned}$$

où $M_\sigma = (\delta_{\sigma^{-1}(i),j})_{i,j} = (\delta_{i,\sigma(j)})_{i,j}$.

(4) Si $n \geqslant 1$, l'hyperplan $H_n = \{(x_i)_i \in K^n \mid x_1 + \cdots + x_n = 0\}$ est stable sous l'action de $\mathfrak{S}_n$ décrite dans le point précédent. Cela fournit une représentation de $\mathfrak{S}_n$ de degré $n-1$.

(5) L'isomorphisme $\mathfrak{S}_3 \simeq D_3 \subset \mathrm{O}(\mathbb{R}^2)$ fournit une représentation de $\mathfrak{S}_3$ de degré 2. Par exemple, si $c = (1\ 2\ 3)$ et $t = (2\ 3)$, il existe un unique morphisme $\rho : \mathfrak{S}_3 \longrightarrow \mathrm{GL}(\mathbb{R}^2)$ tel que

$$\rho(c) = r \quad \text{et} \quad \rho(t) = s,$$

où r est la rotation de centre O et de mesure d'angle $\dfrac{2\pi}{3}$, et s est la symétrie orthogonale d'axe (Ox).

En considérant les matrices représentatives des transformations r et s dans la base canonique de $\mathbb{R}^2$, cela fournit une représentation matri-

cielle $\rho_1 : \mathfrak{S}_3 \longrightarrow \mathrm{GL}_2(\mathbb{R})$, avec

$$\rho_1(c) = \begin{pmatrix} -\frac{1}{2} & -\frac{\sqrt{3}}{2} \\ \frac{\sqrt{3}}{2} & -\frac{1}{2} \end{pmatrix} \quad \text{et} \quad \rho_1(t) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Le point précédent fournit quant à lui une autre représentation H_3 de degré 2. Soient

$$e_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad \text{et} \quad e_2 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$$

Alors, $\mathbf{e} = (e_1, e_2)$ est une base de H_3. On vérifie sans peine les égalités

$$c{\cdot}e_1 = -e_2, \quad c{\cdot}e_2 = e_1 - e_2, \quad t{\cdot}e_1 = e_2 \quad \text{et} \quad t{\cdot}e_2 = e_1.$$

La représentation matricielle correspondante $\rho_2 : \mathfrak{S}_3 \longrightarrow \mathrm{GL}_2(\mathbb{R})$ vérifie donc

$$\rho_2(c) = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} \quad \text{et} \quad \rho_2(t) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

Ces deux représentations sont en fait isomorphes.

En effet, si l'on pose $P = \begin{pmatrix} \sqrt{3} & \sqrt{3} \\ 1 & -1 \end{pmatrix}$, on a

$$P\rho_2(\sigma)P^{-1} = \rho_1(\sigma), \quad \text{pour tout } \sigma \in \mathfrak{S}_3.$$

Notons qu'il suffit de le vérifier pour les deux générateurs c et t de $\mathfrak{S}_3$.

(6) Soit V un K-espace vectoriel. Tout morphisme $\mathbb{Z} \longrightarrow \mathrm{GL}(V)$ est de la forme

$$\begin{aligned} \rho_u : \mathbb{Z} &\longrightarrow \mathrm{GL}(V) \\ m &\longmapsto u^m, \end{aligned}$$

pour un unique automorphisme $u \in \mathrm{GL}(V)$.

On vérifie alors aisément que $\rho_u : \mathbb{Z} \longrightarrow \mathrm{GL}(V)$ et $\rho_{u'} : \mathbb{Z} \longrightarrow \mathrm{GL}(V)$ sont isomorphes si, et seulement si, il existe un isomorphisme

$$f : V \xrightarrow{\sim} V'$$

tel que $f \circ u \circ f^{-1} = u'$. Autrement dit, classer les représentations linéaires de $\mathbb{Z}$ à isomorphisme près revient à savoir classer les automorphismes d'un K-espace vectoriel à similitude près.

(7) Soit $n \geqslant 1$. Les morphismes $\mathbb{Z}/n\mathbb{Z} \longrightarrow \mathrm{GL}(V)$ étant en correspondance bijective avec les morphismes $\rho : \mathbb{Z} \longrightarrow \mathrm{GL}(V)$ tels que $\rho(n) = \mathrm{Id}_V$, les représentations de $\mathbb{Z}/n\mathbb{Z}$ sont toutes de la forme

$$\begin{aligned} \rho_u : \mathbb{Z}/n\mathbb{Z} &\longrightarrow \mathrm{GL}(V) \\ \overline{m} &\longmapsto u^m, \end{aligned}$$

pour un unique automorphisme $u \in \mathrm{GL}(V)$ vérifiant $u^n = \mathrm{Id}_V$.

(8) On suppose que $\mathrm{car}(K) \nmid n$ et que l'ensemble des racines n-ième de l'unité $\mu_n(\overline{K}) \subset K$. Alors, les représentations de degré 1 du groupe $\mathbb{Z}/n\mathbb{Z}$ sont les morphismes

$$\begin{aligned} \mathbb{Z}/n\mathbb{Z} &\longrightarrow K^\times \\ \overline{m} &\longmapsto \zeta^m, \end{aligned}$$

où $\zeta \in \mu_n(\overline{K})$. Elles sont deux à deux non isomorphes. Cela découle immédiatement de l'exemple précédent, ou d'un calcul direct.

Venons-en maintenant aux notions de sous-représentation et de représentation irréductible.

2.20. Définition. Soit V une représentation de G. Une *sous-représentation* de G est un sous-espace W de V stable sous l'action de G. Autrement dit, une sous-représentation de V est un sous-module du $K[G]$-module V.

Enfin, si $\rho : G \longrightarrow \mathrm{GL}(V)$ est le morphisme de groupes correspondant, une sous-représentation W est stable par $\rho(g)$ pour tout $g \in G$. Le morphisme associé à W est alors simplement

$$\begin{aligned} \rho' : G &\longrightarrow \mathrm{GL}(W) \\ g &\longmapsto \rho(g)_W. \end{aligned}$$

On dit qu'une représentation V de G est *irréductible* si V est non nul, et ses seules sous-représentations sont 0 et V. Cela revient à dire que V est un $K[G]$-module simple.

2.21. Remarques. Soit G un groupe.

(1) Si V et V' sont deux représentations de G isomorphes, alors V est irréductible si, et seulement si, V' l'est.

(2) Si G est un groupe fini, une représentation irréductible est nécessairement de degré fini (cf. exercice 4).

2.22. Exemples. Soit G un groupe.

(1) Pour toute représentation V de G, V^G est une sous-représentation de V. Le morphisme correspondant $G \longrightarrow \mathrm{GL}(V^G)$ est le morphisme trivial.

(2) Toute représentation de degré 1 est irréductible, puisque les seuls sous-espaces d'un espace vectoriel V de dimension 1 sont 0 ou V.

En particulier, la représentation triviale est irréductible.

(3) Si V est une représentation de G, pour tout morphisme $\varphi : G \longrightarrow K^\times$, on note $V(\varphi)$ l'espace vectoriel V, muni de l'action

$$\begin{aligned} G \times V &\longrightarrow V \\ (g, v) &\longmapsto g * v = \varphi(g)(g \cdot v). \end{aligned}$$

Si $\rho : G \longrightarrow \mathrm{GL}(V)$ est le morphisme correspondant à l'action de G, le morphisme correspondant à $V(\varphi)$ est donné par

$$\begin{aligned} \varphi \cdot \rho : G &\longrightarrow \mathrm{GL}(V) \\ g &\longmapsto \varphi(g)\rho(g). \end{aligned}$$

On voit facilement qu'un sous-espace vectoriel W de V est stable sous l'action de G si, et seulement si, il est stable sous l'action $*$. En particulier, V est irréductible si, et seulement si, $V(\varphi)$ l'est.

(4) La représentation de $\mathfrak{S}_n$ de degré 1 donnée par la signature est irréductible par (2).

(5) Une représentation de permutation $K[E]$ n'est jamais irréductible, dès que $|E| \geqslant 2$.

Si le groupe G agit trivialement sur E, n'importe quel sous-espace vectoriel de $K[E]$ est G-stable. Il suffit d'en choisir un qui soit distinct de 0 et $K[E]$ pour conclure (ce qui est possible, car $|E| \geqslant 2$).

Supposons à présent que G n'agisse pas trivialement sur E. D'après l'exemple 2.13, $K[E]^G$ est l'ensemble des fonctions constantes sur les G-orbites.

Alors, $K[E]^G$ est un sous-espace G-stable non nul, et distinct de $K[E]$. En effet, par hypothèse, il existe $x \in E$ et $g \in G$ tel que $g \cdot x \neq x$. La fonction e_x est alors un élément de $K[E] \setminus K[E]^G$.

(6) Soit $n \geqslant 2$ un entier. Alors, la représentation H_n de $\mathfrak{S}_n$ définie dans l'exemple 2.19 est une sous-représentation de la représentation K^n définie au point (3) de ce même exemple.

Notons que la représentation H_2 est irréductible, puisqu'elle est de degré 1.

Si $n \geqslant 3$, H_n est irréductible si, et seulement si, $\mathrm{car}(K) \nmid n$.

Supposons tout d'abord que $\mathrm{car}(K) \mid n$. Alors, $v = (1, \ldots, 1) \in H_n$, puisque $n \cdot 1 = 0$ par hypothèse. De plus, v est clairement fixe sous l'action de $\mathfrak{S}_n$. Par conséquent, Kv est une sous-représentation de H_n, distincte de $\{0\}$ et de H_n (puisque $n \geqslant 3$), et H_n n'est pas irréductible.

Supposons maintenant que $\mathrm{car}(K) \nmid n$. Soit V une sous-représentation non nulle de H_n. On doit démontrer que $V = H_n$.

Soit $v = (v_1, \ldots, v_n) \in V \setminus \{0\}$. On commence par remarquer qu'il existe deux entiers $i, j \in [\![1, n]\!]$, $i \neq j$ tels que $v_i \neq v_j$. En effet, dans

le cas contraire, puisque $v \in H_n$, on aurait $n \cdot v_1 = 0$, puis $v_1 = 0$ par hypothèse sur la caractéristique de K, et enfin $v = 0$.

Pour tout $k \in [\![1, n]\!], k \neq i, j$, on a alors

$$\frac{(i\ j\ k)\cdot v - (i\ k)\cdot v}{v_i - v_j} = e_j - e_k \in V,$$

où $(e_1, \ldots, e_n)$ est la base canonique de K^n.

Or, les vecteurs $e_j - e_k$, $k \neq j$, forment une K-base de H_n. On en déduit alors que $V = H_n$. Par conséquent, H_n est bien irréductible dans ce cas.

(7) Si $\mathrm{car}(K) \nmid n$, on montre que les sous-représentations de K^n (muni de l'action naturelle de $\mathfrak{S}_n$) sont

$$0,\ K^n,\ H_n \ \text{ et } \ \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}.$$

Le lecteur se référera à l'exercice 3 pour une démonstration.

(8) Si $\varepsilon : \mathfrak{S}_n \longrightarrow K^\times$ est le morphisme de signature, la représentation linéaire $H_n(\varepsilon)$ de $\mathfrak{S}_n$ est irréductible si $n = 2$, ou si $n \geqslant 3$ et $\mathrm{car}(K) \nmid n$.

Notons qu'il se peut très bien que H_n et $H_n(\varepsilon)$ soient isomorphes. Par exemple, H_3 et $H_3(\varepsilon)$ sont isomorphes si $3 \nmid \mathrm{car}(K)$.

En effet, si $\rho : \mathfrak{S}_3 \longrightarrow \mathrm{GL}_2(K)$ est la représentation matricielle correspondant à H_3 décrite dans l'exemple 2.19 (4), et si l'on pose

$$P = \begin{pmatrix} 1 & 2 \\ -2 & -1 \end{pmatrix},$$

on vérifie sans peine que

$$P\rho(\sigma)P^{-1} = \varepsilon(\sigma)\rho(\sigma), \quad \text{pour tout } \sigma \in \mathfrak{S}_3.$$

En revanche, H_2 et $H_2(\varepsilon)$ ne sont pas isomorphes si $\mathrm{car}(K) \neq 2$, puisque la première correspond à l'action triviale sur K, ce qui n'est pas le cas de la seconde.

Nous aurons dans la suite besoin de la notion de produit tensoriel d'espaces vectoriels. Nous n'insisterons pas trop sur cette notion et nous nous limiterons au strict minimum.

3. Interlude : produit tensoriel

Soient V_1, V_2 deux K-espaces vectoriels. Pour tout $v_1 \in V_1$, et tout $v_2 \in V_2$, on note $\{v_1, v_2\} \in \mathscr{F}(V_1 \times V_2, K)$ la fonction définie par

$$\{v_1, v_2\}(v_1', v_2') = \begin{cases} 1 & \text{si } (v_1', v_2') = (v_1, v_2) \\ 0 & \text{sinon.} \end{cases}$$

Il est facile de voir que la famille $\mathscr{I} = (\{v_1, v_2\})_{(v_1,v_2)\in V_1\times V_2}$ est une famille libre de $\mathscr{F}(V_1 \times V_2, K)$. On note $\mathscr{E} = \mathrm{Vect}_K(\mathscr{I})$. Remarquons que $\mathscr{I}$ est une base de $\mathscr{E}$.

Soit $\mathscr{E}'$ le sous-espace vectoriel de $\mathscr{E}$ engendré par les fonctions

$$\{v_1+v_1', v_2\}-\{v_1, v_2\}-\{v_1', v_2\}, \quad \{v_1, v_2+v_2'\}-\{v_1, v_2\}-\{v_1, v_2+v_2'\},$$
$$\{\lambda v_1, v_2\}-\lambda\{v_1, v_2\}, \quad \{v_1, \lambda v_2\}-\lambda\{v_1, v_2\},$$

pour tous v_1, $v_1' \in V_1$, tous v_2, $v_2' \in V_2$, et tout $\lambda \in K$.

Enfin, soit $V_1 \otimes_K V_2$ l'espace vectoriel quotient $\mathscr{E}/\mathscr{E}'$. On note $v_1 \otimes v_2$ la classe d'équivalence de $\{v_1, v_2\}$.

Le théorème suivant résume les propriétés élémentaires de $V_1 \otimes_K V_2$.

3.1. Théorème. *Soient V_1, V_2, V des K-espaces vectoriels. Alors :*

(1) *l'application*

$$\begin{aligned} \tau : V_1 \times V_2 &\longrightarrow V_1 \otimes_K V_2 \\ (v_1, v_2) &\longmapsto v_1 \otimes v_2 \end{aligned}$$

est K-bilinéaire. De plus, les éléments

$$v_1 \otimes v_2, \quad v_1 \in V_1,\ v_2 \in V_2$$

engendrent $V_1 \otimes_K V_2$.

(2) *pour toute application bilinéaire $b : V_1 \times V_2 \longrightarrow V$, il existe une unique application linéaire $\overline{b} : V_1 \otimes_K V_2 \longrightarrow V$ telle que*

$$\overline{b}(v_1 \otimes v_2) = b(v_1, v_2),$$

pour tout $v_1 \in V_1$ et tout $v_2 \in V_2$;

(3) *si $\mathbf{e} = (e_i)_{i\in I}$ est une base de V_1 et $\mathbf{f} = (f_j)_{j\in J}$ est une base de V_2, alors la famille*

$$\mathbf{e} \otimes \mathbf{f} = (e_i \otimes f_j)_{(i,j)\in I\times J}$$

est une base de $V_1 \otimes_K V_2$.

En particulier, $V_1 \otimes_K V_2$ est de dimension finie si, et seulement si, V_1 et V_2 le sont, et dans ce cas l'on a

$$\dim_K(V_1 \otimes_K V_2) = \dim_K(V_1)\dim_K(V_2).$$

Démonstration. La première partie du point (1) découle du fait que la classe d'un élément de $\mathscr{E}'$ est nulle dans le quotient $V_1 \otimes_K V_2 = \mathscr{E}/\mathscr{E}'$, et de la définition des lois d'un espace vectoriel quotient. De plus, comme la famille $\mathscr{I} = (\{v_1, v_2\})_{(v_1,v_2)\in V_1\times V_2}$ est une base de $\mathscr{E}$, les classes de ses éléments engendrent $V_1 \otimes_K V_2$, d'où la seconde partie.

Soit $b : V_1 \times V_2 \longrightarrow V$ une application bilinéaire.

Comme $\mathscr{I} = (\{v_1, v_2\})_{(v_1,v_2)\in V_1\times V_2}$ est une base de $\mathscr{E}$, il existe une unique application linéaire $f : \mathscr{E} \longrightarrow K$ telle que

$$f(\{v_1, v_2\}) = b(v_1, v_2), \quad \text{pour tout } (v_1, v_2) \in V_1 \times V_2.$$

La bilinéarité de b implique aisément que $\mathscr{E}' \subset \mathrm{Ker}(f)$. Par conséquent, f induit par passage au quotient une application linéaire $\overline{b} : V_1 \otimes_K V_2 \longrightarrow V$ vérifiant

$$\overline{b}(v_1 \otimes v_2) = \overline{b}\big(\overline{\{v_1, v_2\}}\big) = b(v_1, v_2),$$

pour tout $v_1 \in V_1$ et tout $v_2 \in V_2$. Comme la famille $(v_1 \otimes v_2)_{(v_1,v_2)\in V_1\times V_2}$ engendre $V_1 \otimes_K V_2$, une application linéaire $\overline{b}$ vérifiant les conditions de l'énoncé est nécessairement unique, puisque les images des générateurs sont entièrement déterminées, d'où (2).

Montrons (3). Par (1), tout élément de $V_1 \otimes_K V_2$ est combinaison linéaire d'éléments de la forme $v_1 \otimes v_2, v_i \in V_i$. En utilisant la bilinéarité de τ, et le fait que $\mathbf{e}$ et $\mathbf{f}$ sont des bases de V_1 et V_2 respectivement, on constate que la famille $\mathbf{e} \otimes \mathbf{f}$ engendre $V_1 \otimes_K V_2$. Montrons que c'est une famille libre.

Soit $(\lambda_{ij})_{(i,j)\in I\times J}$ une famille d'éléments de K à support fini, telle que

$$\sum_{i,j} \lambda_{ij} e_i \otimes f_j = 0.$$

Soit $(k, \ell) \in I \times J$, et soit $b_{k,\ell} : V_1 \times V_2 \longrightarrow K$ l'unique application bilinéaire telle que

$$b_{k,\ell}(e_i, f_j) = \delta_{i,k}\delta_{j,\ell}, \quad \text{pour tout } (i, j) \in I \times J.$$

D'après le point précédent, il existe alors une unique application linéaire

$$f_{k,\ell} : V_1 \otimes_K V_2 \longrightarrow K$$

telle que

$$f_{k,\ell}(v_1 \otimes v_2) = b_{k,\ell}(v_1, v_2), \quad \text{pour tout } v_i \in V_i.$$

En particulier, on a

$$f_{k,\ell}(e_i \otimes f_j) = \delta_{i,k}\delta_{j,\ell}, \quad \text{pour tout } (i, j) \in I \times J.$$

En appliquant $f_{k,\ell}$ à la relation de dépendance linéaire ci-dessus, on obtient alors $\lambda_{k\ell} = 0$.

Ainsi, la famille $\mathbf{e}\otimes\mathbf{f}$ est une base de $V_1 \otimes_K V_2$, et le reste en découle immédiatement. □

3.2. Définition. L'espace vectoriel $V_1 \otimes_K V_2$ s'appelle le *produit tensoriel* de V_1 et V_2.

Le lemme suivant nous sera utile pour définir le produit de deux représentations.

3.3. Lemme. *Soient V_1, V_2, V_1', V_2' des K-espaces vectoriels. Pour toutes applications linéaires $f_1 : V_1 \longrightarrow V_1'$ et $f_2 : V_2 \longrightarrow V_2'$, il existe une unique application linéaire $f_1 \otimes f_2 : V_1 \otimes_K V_2 \longrightarrow V_1' \otimes_K V_2'$ telle que*

$$(f_1 \otimes f_2)(v_1 \otimes v_2) = f_1(v_1) \otimes f_2(v_2), \quad \text{pour tout } v_i \in V_i.$$

Démonstration. La linéarité de f_1 et f_2 et la bilinéarité de l'application canonique $V_1' \times V_2' \longrightarrow V_1' \otimes_K V_2'$ impliquent que l'application

$$\begin{aligned} V_1 \times V_2 &\longrightarrow V_1' \otimes_K V_2' \\ (v_1, v_2) &\longmapsto f_1(v_1) \otimes f_2(v_2) \end{aligned}$$

est bilinéaire. Il suffit alors d'appliquer le point (2) du théorème 3.1 pour conclure. □

Soit L/K une extension de corps. Remarquons qu'une K-base $\mathbf{e}$ de K^n reste une L-base de L^n, lorsqu'on la regarde comme une famille de vecteurs de L^n. En effet, si P est la matrice dont les colonnes sont les éléments de la base $\mathbf{e}$, alors $P \in \mathrm{GL}_n(K)$. Mais alors, $P \in \mathrm{GL}_n(L)$ et $\mathbf{e}$ est une L-base de L^n. Nous allons maintenant généraliser ces considérations.

Pour tout $\lambda \in L$, l'application

$$\begin{aligned} V \times L &\longrightarrow V \otimes_K L \\ (v, \lambda') &\longmapsto v \otimes \lambda\lambda' \end{aligned}$$

est K-bilinéaire. Il existe donc une unique application K-linéaire

$$f_\lambda : V \otimes_K L \longrightarrow V \otimes_K L$$

telle que

$$f_\lambda(v \otimes \lambda') = v \otimes \lambda\lambda', \quad \text{pour tout } v \in V \text{ et tout } \lambda' \in L.$$

Dans la suite, pour tout $\lambda \in L$ et tout $z \in V \otimes_K L$, on pose $\lambda \bullet z = f_\lambda(z)$. On a donc

$$\lambda \bullet (v \otimes \lambda') = v \otimes \lambda\lambda', \quad \text{pour tout } v \in V \text{ et tout } \lambda' \in L.$$

Si $(e_i)_{i \in I}$ est une K-base de V et $(f_j)_{j \in J}$ est une K-base de L, on a ainsi

$$\lambda \bullet \Big(\sum_{i,j} \lambda_{ij} (e_i \otimes f_j) \Big) = \sum_{i,j} \lambda_{ij} (e_i \otimes \lambda f_j), \quad \text{pour tout } \lambda_{ij} \in K.$$

3.4. Proposition. *Soit V un K-espace vectoriel et soit L/K une extension. Alors, la loi externe*

$$\begin{aligned} L \times (V \otimes_K L) &\longrightarrow V \otimes_K L \\ (\lambda, z) &\longmapsto \lambda \bullet z \end{aligned}$$

munit le groupe abélien $V \otimes_K L$ d'une structure de L-espace vectoriel.

De plus, pour tout $\lambda \in K$, et tout $z \in V \otimes_K L$, on a $\lambda \bullet z = \lambda z$.

Enfin, si $\mathbf{e} = (e_i)_{i\in I}$ est une K-base de V, alors $\mathbf{e} \otimes \mathbf{1} = (e_i \otimes 1)_{i\in I}$ est une L-base de $V \otimes_K L$. En particulier, $V \otimes_K L$ est de dimension finie sur L si, et seulement si, V est de dimension finie, et dans ce cas, on a

$$\dim_L(V \otimes_K L) = \dim_K(V).$$

Démonstration. Le fait que $V \otimes_K L$ soit un L-espace vectoriel résulte de calculs élémentaires laissés au lecteur.

Soit $\mathbf{e} = (e_i)_{i\in I}$ une K-base de V. Soit $\mathbf{f} = (f_j)_{j\in J}$ une K-base de L. Par le théorème 3.1, la famille $(e_i \otimes f_j)_{(i,j)\in I\times J}$ est une K-base de $V \otimes_K L$. Pour tous λ, $\lambda_{ij} \in K$, on a alors

$$\lambda \bullet \Big(\sum_{i,j} \lambda_{ij}(e_i \otimes f_j)\Big) = \sum_{i,j} \lambda_{ij}(e_i \otimes \lambda f_j) = \lambda\Big(\sum_{i,j} \lambda_{ij}(e_i \otimes f_j)\Big),$$

la dernière égalité provenant de la K-bilinéarité de $V \times L \longrightarrow V \otimes_K L$. Cela démontre la deuxième assertion.

Par ailleurs, on a

$$e_i \otimes f_j = e_i \otimes 1 \cdot f_j = f_j \bullet (e_i \otimes 1), \qquad \text{pour tout } (i,j) \in I \times J,$$

ce qui implique aisément que $\mathbf{e} \otimes 1$ est L-génératrice.

Soit $(\lambda_i)_{i\in I}$ une famille d'éléments de L à support fini telle que

$$\sum_i \lambda_i \bullet (e_i \otimes 1) = 0.$$

Écrivons $\lambda_i = \sum_j \lambda_{ij} f_j$, avec $\lambda_{ij} \in K$. On a alors

$$\sum_i \lambda_i \bullet (e_i \otimes 1) = \sum_i e_i \otimes \lambda_i = \sum_{i,j} \lambda_{ij} e_i \otimes f_j = 0.$$

Puisque $(e_i \otimes f_j)_{(i,j)\in I\times J}$ est une K-base de $V \otimes_K L$, on en déduit que l'on a $\lambda_{ij} = 0$ pour tout $i \in I$ et tout $j \in J$. Par conséquent, $\lambda_i = 0$ pour tout $i \in I$. Ainsi, $\mathbf{e} \otimes 1$ est bien une L-base de $V \otimes_K L$.

Le reste de la proposition est clair. □

On termine cette section par un lemme.

3.5. Lemme. *Soient V, V' deux K-espaces vectoriels et soit L/K une extension. Pour toute application linéaire $f : V \longrightarrow V'$, il existe une unique application L-linéaire $f_L : V \otimes_K L \longrightarrow V' \otimes_K L$ telle que*

$$f_L(v \otimes 1) = f(v) \otimes 1 \quad \textit{pour tout } v \in V.$$

On a $(\mathrm{Id}_V)_L = \mathrm{Id}_{V \otimes_K L}$, et si $g : V' \longrightarrow V''$ est K-linéaire, on a

$$(g \circ f)_L = g_L \circ f_L.$$

En particulier, si f est un isomorphisme, alors f_L est un isomorphisme, d'inverse $(f^{-1})_L$.

Si de plus les K-espaces vectoriels V et V' sont de dimension finie, et si $\mathbf{e}$ et $\mathbf{e}'$ sont des bases de V et de V' respectivement, on a

$$\mathrm{Mat}(f_L\,; \mathbf{e} \otimes 1, \mathbf{e}' \otimes 1) = \mathrm{Mat}(f\,; \mathbf{e}, \mathbf{e}').$$

Démonstration. Soit $\mathbf{e} = (e_i)_{i \in I}$ une K-base de V. Si une application f_L satisfaisant les conditions de l'énoncé existe, on a

$$f_L(e_i \otimes 1) = f(e_i) \otimes 1, \quad \text{pour tout } i \in I.$$

D'après la proposition précédente, $\mathbf{e} \otimes 1$ est une L-base de $V \otimes_K L$. Les relations ci-dessus montrent alors que f_L est unique.

Réciproquement, soit $f_L : V \otimes_K L \longrightarrow V' \otimes_K L$ l'application $f \otimes \mathrm{Id}_L$. Cette application est K-linéaire, et vérifie

$$f_L(v \otimes 1) = f(v) \otimes 1, \quad \text{pour tout } v \in V.$$

Il reste donc à démontrer que f_L est L-linéaire. Pour tout $\lambda_i \in L$, on a

$$f_L\Big(\sum_i \lambda_i \bullet (e_i \otimes 1)\Big) = f_L\Big(\sum_i e_i \otimes \lambda_i\Big) = \sum_i f(e_i) \otimes \lambda_i.$$

Mais alors, pour tout $\lambda \in L$, on a

$$\begin{aligned} f_L\Big(\lambda \bullet \Big(\sum_i \lambda_i \bullet (e_i \otimes 1)\Big)\Big) &= f_L\Big(\sum_i e_i \otimes \lambda\lambda_i\Big) \\ &= \sum_i f(e_i) \otimes \lambda\lambda_i \\ &= \lambda \bullet \Big(\sum_i f(e_i) \otimes \lambda_i\Big) \\ &= \lambda \bullet f_L\Big(\sum_i \lambda_i \bullet (e_i \otimes 1)\Big). \end{aligned}$$

Les égalités $(\mathrm{Id}_V)_L = \mathrm{Id}_{V \otimes_K L}$ et $(g \circ f)_L = g_L \circ f_L$ se vérifient alors immédiatement en regardant les images des éléments de la base $\mathbf{e} \otimes 1$.

Enfin, supposons que V et V' soient de dimension finie. Soit $\mathbf{e} = (e_1, \ldots, e_n)$ une K-base de V, et soit $\mathbf{e}' = (e'_1, \ldots, e'_m)$ une K-base de V'.

Si $\mathrm{Mat}(f\,; \mathbf{e}, \mathbf{e}') = (a_{ij})_{i,j} \in \mathrm{M}_{m\times n}(K)$, on a

$$f(e_j) = \sum_{i=1}^{n} a_{ij} e'_i, \quad \text{pour tout } j \in [\![1, n]\!].$$

Mais alors, on a

$$f_L(e_j \otimes 1) = f(e_j) \otimes 1 = \sum_{i=1}^{n} a_{ij}(e'_i \otimes 1) = \sum_{i=1}^{n} a_{ij} \bullet (e'_i \otimes 1),$$

pour tout $j \in [\![1, n]\!]$, d'où l'égalité

$$\mathrm{Mat}(f_L\,; \mathbf{e} \otimes 1, \mathbf{e}' \otimes 1) = \mathrm{Mat}(f\,; \mathbf{e}, \mathbf{e}').$$

Cela achève la démonstration. □

Nous allons maintenant exposer diverses manières de construire de nouvelles représentations.

4. Opérations sur les représentations

Commençons par la notion de somme directe externe.

Si $V_1, \ldots V_r$ sont des représentations de G, on définit une action de G sur l'espace vectoriel $V_1 \times \cdots \times V_r$ par

$$\begin{aligned} G \times (V_1 \times \cdots \times V_r) &\longrightarrow V_1 \times \cdots \times V_r \\ \big(g, (v_1, \ldots, v_r)\big) &\longmapsto (g{\cdot}v_1, \ldots, g{\cdot}v_r). \end{aligned}$$

Cette action étant linéaire, on obtient une représentation de G.

En termes de $K[G]$-modules, c'est simplement la somme directe externe des modules $V_1, \ldots, V_r$.

Si $\rho_i : G \longrightarrow \mathrm{GL}(V_i)$ est le morphisme correspondant à la représentation linéaire V_i, le morphisme associé à la représentation $V_1 \times \cdots \times V_r$ est le morphisme $\rho : G \longrightarrow \mathrm{GL}(V_1 \times \cdots \times V_r)$ défini par

$$\rho(g)(v_1, \ldots, v_r) = \big(\rho_1(g)(v_1), \ldots, \rho_r(g)(v_r)\big),$$

pour tout $g \in G$ et tout $v_i \in V_i$.

Ce morphisme sera noté $\rho_1 \oplus \cdots \oplus \rho_r$.

Supposons que $V_1, \ldots, V_r$ soient de dimension finie. Le choix d'une base $\mathbf{e}_i$ de V_i donne lieu à une représentation matricielle

$$\rho'_i : G \longrightarrow \mathrm{M}_{n_i}(K)$$

isomorphe à ρ_i.

En utilisant la base de $V_1 \times \cdots \times V_r$ construite de manière évidente à partir de celles de $V_1, \ldots, V_r$, on voit que ρ est isomorphe à la représentation matricielle

$$\begin{aligned} \rho' : G &\longrightarrow \quad \mathrm{M}_n(K) \\ g &\longmapsto \begin{pmatrix} \rho'_1(g) & & \\ & \ddots & \\ & & \rho'_r(g) \end{pmatrix}, \end{aligned}$$

où $n = \dim_K(V_1) + \cdots + \dim_K(V_r)$.

4.1. Définition. La représentation $V_1 \times \cdots \times V_r$ est appelée la *somme directe externe* des représentations $V_1, \ldots, V_r$.

On continue avec la notion de somme directe interne.

4.2. Définition. On dit qu'une représentation V de G sur K est la *somme directe interne* des sous-représentations $V_1, \ldots, V_r$ si $V = V_1 \oplus \cdots \oplus V_r$ comme K-espace vectoriel.

Cela revient au même de dire que le $K[G]$-module V est somme directe interne des sous-modules $V_1, \ldots, V_r$.

En choisissant une base $\mathbf{e}_i$ de V_i et en recollant ces bases pour obtenir une base $\mathbf{e}$ de V, on voit que pour tout $g \in G$, on a

$$\mathrm{Mat}\big(\rho(g)\,;\mathbf{e}\big) = \begin{pmatrix} \mathrm{Mat}\big(\rho(g)_{V_1}\,;\mathbf{e}_1\big) & & \\ & \ddots & \\ & & \mathrm{Mat}\big(\rho(g)_{V_r}\,;\mathbf{e}_r\big) \end{pmatrix}.$$

Dans ce cas, on a bien entendu un isomorphisme de représentations

$$(V,\rho) \simeq (V_1,\rho_1) \times \cdots \times (V_r,\rho_r), \quad \text{où} \quad \begin{aligned} \rho_i : G &\longrightarrow \mathrm{GL}(V_i) \\ g &\longmapsto \rho(g)_{V_i}. \end{aligned}$$

4.3. Exemples

(1) Soit V un K-espace vectoriel de dimension finie $n \geqslant 1$, sur lequel G agit trivialement. Alors, V est isomorphe à la somme directe externe de n copies de la représentation triviale $\mathbf{1}$.

(2) Soit G un groupe fini, et soit K un corps tel que $\mathrm{car}(K) \nmid |G|$. Pour toute représentation V, on a

$$V = V^G \oplus V_G,$$

où

$$V_G = \Big\{ v \in V \,\Big|\, \sum_{g \in G} g \cdot v = 0 \Big\}.$$

Cela découle directement du fait que p_G est un projecteur, où p_G est défini dans le lemme 2.5.

(3) Soit $n \geqslant 2$, et soit K un corps tel que $\mathrm{car}(K) \nmid n!$. Considérons l'action de $\mathfrak{S}_n$ sur K^n par permutation des coordonnées. L'exemple précédent montre que l'on a

$$K^n = K \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \oplus H_n.$$

Notons que c'est la seule décomposition non triviale de cette représentation que l'on puisse obtenir, puisque les sous-représentations de K^n sont 0, K^n, H_n et $K \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$, d'après l'exemple 2.22 (7).

4.4. Exemple. Soit V une représentation linéaire de G. Pour toute forme linéaire $\varphi \in V^*$, on pose

$$(g{\cdot}\varphi)(v) = \varphi(g^{-1}{\cdot}v), \quad \text{pour tout } g \in G \text{ et tout } v \in V.$$

On vérifie aisément que $g{\cdot}\varphi \in V^*$, et que l'application

$$\begin{aligned} G \times V^* &\longrightarrow V^* \\ (g, \varphi) &\longmapsto g{\cdot}\varphi \end{aligned}$$

définit une action linéaire de G sur V^*, i.e. une représentation de G.

Si $\rho : G \longrightarrow \mathrm{GL}(V)$ est le morphisme correspondant à l'action de G, le morphisme correspondant à la représentation précédente est donné par

$$\begin{aligned} G &\longrightarrow \mathrm{GL}(V^*) \\ g &\longmapsto \rho(g^{-1})^*. \end{aligned}$$

4.5. Définition. Soit V une représentation linéaire de G. La représentation linéaire V^* définie dans l'exemple précédent s'appelle la *représentation duale* de V.

Passons maintenant au produit tensoriel de représentations.

4.6. Définition. Soient maintenant (V_1, ρ_1) et (V_2, ρ_2) deux représentations de G. Le *produit tensoriel* est la représentation $(V_1 \otimes_K V_2,\ \rho_1 \otimes \rho_2)$, où $\rho_1 \otimes \rho_2 : G \longrightarrow \mathrm{GL}(V_1 \otimes_K V_2)$ est défini par

$$(\rho_1 \otimes \rho_2)(g) = \rho_1(g) \otimes \rho_2(g), \quad \text{pour tout } g \in G.$$

On vérifie aisément que l'on obtient bien une représentation de G.

Elle correspond à l'unique action de G sur $V_1 \otimes_K V_2$ telle que

$$g{\cdot}(v_1 \otimes v_2) = (g{\cdot}v_1) \otimes (g{\cdot}v_2),$$

pour tout $g \in G$, et tout $v_i \in V_i$.

4.7. Remarque. Supposons que V_1 et V_2 soient de dimension finie.

Soit $\mathbf{e} = (e_1, \ldots, e_n)$ une base de V_1 et soit $\mathbf{f} = (f_1, \ldots, f_m)$ une base de V_2. Soit $g \in G$. Posons $A = \mathrm{Mat}\big(\rho_1(g)\,;\mathbf{e}\big) = (a_{ij})_{i,j}$ et $B = \mathrm{Mat}\big(\rho_2(g)\,;\mathbf{f}\big)$. Alors, en ordonnant les éléments de $[\![1,n]\!] \times [\![1,m]\!]$ grâce à l'ordre lexicographique, on obtient

$$\mathrm{Mat}\Big(\big(\rho_1(g) \otimes \rho_2(g)\big)\,;\mathbf{e}\otimes\mathbf{f}\Big) = (a_{k\ell}B)_{k,\ell} \in \mathrm{M}_{nm}(K).$$

Lorsque les représentations V_1 et V_2 sont égales à une même représentation V, on peut définir deux sous-représentations intéressantes de $V \otimes_K V$.

Notons tout d'abord que l'application

$$\begin{aligned} V \times V &\longrightarrow V \otimes_K V \\ (v, v') &\longmapsto sv' \otimes v \end{aligned}$$

est bilinéaire, et induit donc une application linéaire $s_V \in \mathscr{L}(V \otimes_K V)$ telle que

$$s_V(v \otimes v') = v' \otimes v, \quad \text{pour tous } v,\, v' \in V.$$

Pour tout $g \in G$, et tous v, $v' \in V$, on a

$$\begin{aligned} s_V\big(g\cdot(v \otimes v')\big) &= s_V\big((g\cdot v) \otimes (g\cdot v')\big) \\ &= (g\cdot v') \otimes (g\cdot v) \\ &= g\cdot(v' \otimes v) \\ &= g\cdot s_V(v \otimes v'). \end{aligned}$$

Comme la famille $(v \otimes v')_{v,v' \in V}$ engendre $V \otimes_K V$, la linéarité de l'action de G montre que

$$s_V(g\cdot x) = g\cdot s_V(x), \quad \text{pour tout } g \in G \text{ et tout } x \in V \otimes_K V.$$

Il s'ensuit que les sous-espaces

$$\mathrm{Sym}^2(V) = \{x \in V \otimes_K V \mid s_V(x) = x\}$$

et

$$\mathrm{Alt}^2(V) = \{x \in V \otimes_K V \mid s_V(x) = -x\}$$

sont G-stables.

4.8. Définition. Soit V une représentation linéaire de G.

Les sous-représentations $\mathrm{Sym}^2(V)$ et $\mathrm{Alt}^2(V)$ s'appellent respectivement le *carré symétrique* et le *carré alterné* de V.

4.9. Remarque. Soit V une représentation de G. On a $s_V^2 = \mathrm{Id}_{V\otimes_K V}$, comme on peut le vérifier sur la famille génératrice $(v \otimes v')_{v,v' \in V}$. En particulier, si $\mathrm{car}(K) \neq 2$, on a

$$V \otimes_K V = \mathrm{Sym}^2(V) \oplus \mathrm{Alt}^2(V).$$

Dans ce cas, si $\dim_K(V) = n$, et si $\mathbf{e} = (e_i)_{i\in[\![1,n]\!]}$ est une base de V, les familles

$$\mathrm{Sym}^2(\mathbf{e}) = (e_i \otimes e_j + e_j \otimes e_i)_{i\leqslant j} \quad \text{et} \quad \mathrm{Alt}^2(V) = (e_i \otimes e_j - e_j \otimes e_i)_{i<j}$$

sont des bases de $\mathrm{Sym}^2(V)$ et $\mathrm{Alt}^2(V)$ respectivement. On a donc

$$\dim_K\left(\mathrm{Sym}^2(V)\right) = \frac{n(n+1)}{2} \quad \text{et} \quad \dim_K\left(\mathrm{Alt}^2(V)\right) = \frac{n(n-1)}{2}.$$

Continuons en définissant l'extension des scalaires d'une représentation.

Soit V une représentation de G sur un corps K et soit L/K une extension de corps. Si $\rho : V \longrightarrow \mathrm{GL}(V)$ est le morphisme correspondant, on considère le morphisme

$$\begin{aligned} \rho_L : G &\longrightarrow \mathrm{GL}(V \otimes_K L) \\ g &\longmapsto \quad \rho(g)_L. \end{aligned}$$

Ce morphisme correspond à l'unique action L-linéaire de G sur $V \otimes_K L$ telle que

$$g\cdot(v \otimes 1) = (g\cdot v) \otimes 1, \qquad \text{pour tout } g \in G \text{ et tout } v \in V.$$

4.10. Définition. La représentation $V \otimes_K L$ s'appelle *l'extension des scalaires de V à L.*

Nous achevons cette section en montrant comment construire une représentation d'un sous-groupe et d'un quotient de G à partir d'une représentation de G.

4.11. Exemple. Soit V une représentation de G et soit H un sous-groupe de G. La restriction de l'action de G à H étant encore linéaire, V est aussi une représentation de H, que l'on note $\mathrm{Res}^G_H(V)$.

Si $\rho : G \longrightarrow \mathrm{GL}(V)$ est le morphisme correspondant à V, le morphisme correspondant à $\mathrm{Res}^G_H(V)$ est simplement

$$\rho_{|H} : H \longrightarrow \mathrm{GL}(V).$$

En termes de modules, cela revient à regarder V comme un $K[H]$-module via l'action

$$\begin{aligned} K[H] \times V &\longrightarrow V \\ \Big(\sum_{h\in H} \lambda_h h, v\Big) &\longmapsto \sum_{h\in H} \lambda_h h\cdot v. \end{aligned}$$

Si $\mathrm{Res}^G_H(V)$ est irréductible, V est alors irréductible. En effet, un sous-espace de V stable sous l'action de G est en particulier stable sous l'action de H.

4.12. Définition. La représentation linéaire $\mathrm{Res}_H^G(V)$ est appelée la *restriction* de G à H de la représentation V.

On considère maintenant le cas des groupes quotients.

4.13. Exemple. Soit $\pi : G \longrightarrow G'$ un morphisme de groupes. Si V' est une représentation linéaire de G', on peut définir une action de G sur V' par

$$\begin{aligned} G \times V' &\longrightarrow V' \\ (g, v') &\longmapsto \pi(g)\cdot v'. \end{aligned}$$

Cette action est encore linéaire, et l'on obtient une représentation linéaire de G, que l'on note $\pi^*(V')$. Notons que par définition, $\mathrm{Ker}(\pi)$ agit trivialement sur $\pi^*(V')$.

Si $\rho' : G' \longrightarrow \mathrm{GL}(V')$ est le morphisme correspondant à V', le morphisme correspondant à $\pi^*(V')$ est

$$\rho' \circ \pi : G \longrightarrow \mathrm{GL}(V).$$

Supposons maintenant que $\pi : G \longrightarrow G'$ soit surjectif. On a alors une correspondance bijective entre l'ensemble des représentations de G' et l'ensemble des représentations de G sur lesquelles $\mathrm{Ker}(\pi)$ agit trivialement.

En effet, soit V est une représentation de G sur laquelle $\mathrm{Ker}(\pi)$ agit trivialement. Soit $g' \in G$ et soient $g_1, g_2 \in G$ tels que $\pi(g_1) = \pi(g_2) = g'$. On a alors $g_2 g_1^{-1} \in \mathrm{Ker}(\pi)$, et par conséquent, pour tout $v \in V$, on a

$$g_2 \cdot v = (g_2 g_1^{-1} g_1)\cdot v = (g_2 g_1^{-1})\cdot(g_1 \cdot v) = g_1 \cdot v.$$

Autrement dit, si $\widetilde{g'} \in G$ est un antécédent de $g' \in G'$ par π, l'élément $\widetilde{g'}\cdot v$ est indépendant du choix de $\widetilde{g'}$. L'application

$$\begin{aligned} G' \times V &\longrightarrow V \\ (g', v) &\longmapsto \widetilde{g'}\cdot v \end{aligned}$$

est donc bien définie. On vérifie que c'est une action linéaire sur V. On obtient ainsi une représentation $\overline{V}$ de G'.

On vérifie alors que les applications

$$\begin{array}{ccc} V' & \longmapsto & \pi^*(V') \\ \overline{V} & \longleftarrow & V \end{array}$$

sont inverses l'une de l'autre.

Enfin, si V' est une représentation irréductible de G', $\pi^*(V')$ est aussi irréductible. En particulier, l'ensemble des représentations irréductibles de G' s'injecte dans l'ensemble des représentations irréductibles de G.

En effet, si W est un sous-espace stable de $\pi^*(V')$, on vérifie que $\overline{W}$ est un sous-espace stable de V' sous l'action de G'. Par hypothèse, on a $\overline{W} = 0$ ou V'. Grâce à la bijection précédente, on obtient que $W = 0$ ou $\pi^*(V')$. Les détails sont laissés au lecteur.

4.14. Définition. Soit $\pi : G \longrightarrow G'$ un morphisme de groupes. Si V' est une représentation de G', la représentation $\pi^*(V')$ est appelée *l'inflation* de G' à G de la représentation V'.

Si π est surjectif, et V est une représentation de G où $\mathrm{Ker}(\pi)$ agit trivialement, la représentation $\overline{V}$ de G' s'appelle la *représentation quotient* de V.

4.15. Remarques

(1) Soit (V, ρ) une représentation de G. Alors, $\mathrm{Ker}(\rho)$ agit trivialement sur V, et V est donc l'inflation de la représentation

$$\overline{\rho} : G/\mathrm{Ker}(\rho) \longrightarrow \mathrm{GL}(V),$$

l'inflation étant relative au morphisme $G \longrightarrow G/\mathrm{Ker}(\rho)$. Ainsi, toute représentation de G est l'inflation d'une représentation fidèle d'un quotient de G.

(2) Puisque $K^\times$ est commutatif, tout morphisme de groupes $\varphi : G \longrightarrow K^\times$ vérifie $[G, G] \subset \mathrm{Ker}(\varphi)$. On a donc une correspondance bijective entre les morphismes $G \longrightarrow K^\times$ et les morphismes $G/[G, G] \longrightarrow K^\times$.

En particulier, tout caractère linéaire de degré 1 de G est l'inflation d'un unique caractère linéaire de degré 1 de $G/[G, G]$. Le groupe quotient $G/[G, G]$ étant abélien, cette observation facilite grandement la détermination des caractères linéaires de G. Cela montre aussi que, lorsque G est fini et $K = \mathbb{C}$, le nombre de représentations irréductibles de degré 1 est égal à l'indice de $[G, G]$ dans le groupe G, cela d'après le corollaire XII-3.6.

5. Théorème de Maschke

Il est bien connu que tout sous-espace vectoriel W d'un espace vectoriel V possède un supplémentaire. Une question naturelle dans notre contexte est la suivante : tout sous-espace G-stable W d'une représentation V admet-elle un supplémentaire G-stable ? En termes de $K[G]$-modules, cela revient à savoir si tout sous-module d'un $K[G]$-module possède un sous-module supplémentaire.

Le théorème de Maschke répond de manière positive à cette question dans le cas des groupes finis, et sous une hypothèse de bonne caractéristique.

5.1. Théorème (Maschke). *Soit G un groupe fini, et soit K un corps tel que* $\mathrm{car}(K) \nmid |G|$. *Alors, tout sous-espace G-stable d'une représentation admet un sous-espace supplémentaire G-stable.*

Autrement dit, tout sous-module W d'un $K[G]$-module V admet un sous-module supplémentaire.

En particulier :

(1) *un $K[G]$-module est indécomposable si, et seulement si, il est simple ;*

(2) *toute représentation V de G de degré fini non nul est somme directe interne d'un nombre fini de sous-représentations irréductibles.*

Démonstration. Soit V une représentation linéaire de G, et soit W un sous-espace G-stable de V. Alors, le sous-espace vectoriel W admet donc un supplémentaire W', même si V est de dimension infinie. Soit $\pi_W \in \mathscr{L}(V)$ la projection sur W parallèlement à W'. Soit $p_W \in \mathscr{L}(V)$ l'application linéaire définie par

$$p_W(v) = \frac{1}{|G|} \sum_{g \in G} g{\cdot}\big(\pi_W(g^{-1}{\cdot}v)\big), \qquad \text{pour tout } v \in V.$$

Nous allons démontrer que p_W est un projecteur sur W vérifiant

$$p_W(g{\cdot}v) = g{\cdot}p_W(v), \qquad \text{pour tout } g \in G \text{ et tout } v \in V.$$

Par définition, $\pi_W(V) = W$. Comme W est un sous-espace G-stable, on en déduit que $p_W(V) \subset W$. Inversement, pour tout vecteur $w \in W$, et pour tout $g \in G$, on a $g^{-1}{\cdot}w \in W$, car W est G-stable, et par conséquent

$$\pi_W(g^{-1}{\cdot}w) = g^{-1}{\cdot}w, \qquad \text{pour tout } g \in G \text{ et tout } w \in W.$$

Il s'ensuit que $w = p_W(w) \in p_W(V)$ pour tout $w \in W$. Ainsi, $p_W(V) = W$. De plus, le fait que la restriction de p_W à W soit l'identité sur W entraîne que $p_W^2 = p_W$. Autrement dit, p_W est un projecteur sur W.

Enfin, pour tout $v \in V$ et tout $g \in G$, on a

$$p_W(g{\cdot}v) = \frac{1}{|G|} \sum_{h \in G} h{\cdot}\big(\pi_W\big(h^{-1}{\cdot}(g{\cdot}v)\big)\big) = \frac{1}{|G|} \sum_{h \in G} h{\cdot}\big(\pi_W\big((h^{-1}g){\cdot}v\big)\big).$$

En faisant le changement de variables $h \longleftrightarrow gh$, on obtient

$$p_W(g{\cdot}v) = \frac{1}{|G|} \sum_{h \in G} gh{\cdot}\big(\pi_W(h^{-1}{\cdot}v)\big) = \frac{1}{|G|} \sum_{h \in G} g{\cdot}\big(h{\cdot}\big(\pi_W(h^{-1}{\cdot}v)\big)\big).$$

En utilisant la linéarité de l'action de G, on obtient finalement

$$p_W(g{\cdot}v) = g{\cdot}p_W(v), \qquad \text{pour tout } g \in G \text{ et tout } v \in V.$$

Puisque p_W est un projecteur sur W, on a

$$V = \mathrm{Im}(p_W) \oplus \mathrm{Ker}(p_W) = W \oplus \mathrm{Ker}(p_W).$$

Or, l'égalité précédente implique que $\mathrm{Ker}(p_W)$ est G-stable, d'où la première partie du théorème.

Un $K[G]$-module simple est bien sûr indécomposable. Inversement, soit V un $K[G]$-module indécomposable. Soit W un sous-module de V. D'après le point précédent, il existe un sous-module W' de V tel que $V = W \oplus W'$. Puisque V est indécomposable, on obtient $W = 0$ ou $W' = 0$, auquel cas on a $W = V$. Ainsi, V est un $K[G]$-module simple, d'où (1).

Supposons enfin que V soit une représentation de G de degré $n \geqslant 1$. Montrons (2) par récurrence sur n. Soit (H_n) la propriété suivante :

(H_n) toute représentation de G de degré $k \in [\![1, n]\!]$ est une somme directe d'un nombre fini de sous-représentations irréductibles.

Puisque toute représentation de degré 1 est irréductible, (H_1) est vraie. Supposons que (H_n) soit vraie pour un certain entier $n \geqslant 1$, et montrons que (H_{n+1}) est vraie. Soit V une représentation de degré $k \in [\![1, n+1]\!]$. Si V est une représentation irréductible, on a fini. Sinon, il existe un sous-espace W G-stable distinct de 0 et V. D'après le premier point, il existe alors un sous-espace W' G-stable tel que $V = W \oplus W'$. Comme W est distinct de 0 et V, il en est de même de W'. En particulier, W et W' sont non nuls, et de dimension inférieure ou égale à n. Par hypothèse de récurrence, W est somme directe interne d'un nombre fini de sous-représentations irréductibles, et de même pour W'. Mais, toute sous-représentation irréductible de W et W' est une sous-représentation irréductible de V.

Comme $V = W \oplus W'$, on en déduit que V est somme directe interne d'un nombre fini de sous-représentations irréductibles. Cela achève la récurrence, ainsi que la démonstration. □

5.2. Corollaire. *Soit G un groupe fini non trivial, et soit K un corps tel que $\mathrm{car}(K) \nmid |G|$. Alors, il existe une représentation irréductible V sur laquelle G agit non trivialement.*

Démonstration. Supposons que l'action de G sur toute représentation irréductible soit triviale. Le théorème de Maschke entraîne alors que G agit trivialement sur toute représentation de degré fini. En particulier, c'est le cas pour la représentation régulière de G. Autrement dit, l'action de G sur lui-même par translation à gauche est l'action triviale. Cela implique que, pour tout $g \in G$, on a

$$1_G = g \cdot 1_G = g.$$

Ainsi, G est trivial. Cela achève la démonstration. □

Remarquons que le théorème de Maschke devient faux si le groupe G est infini ou si $\mathrm{car}(K) \mid |G|$, comme le montre l'exemple suivant.

5.3. Exemple. Soit K un corps, et soit $G = K$. On vérifie que l'on définit une action linéaire de G sur K^2 en posant

$$\lambda \cdot \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x + \lambda y \\ y \end{pmatrix},$$

pour tous λ, x, $y \in K$.

On constate alors facilement que les seules sous-représentations de K^2 sont

$$0, \quad K^2 \quad \text{et} \quad K \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

En particulier, $W = K \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ est une sous-représentation de K^2 sans supplémentaire G-stable. Cela implique de plus que K^2 est indécomposable, mais non irréductible. Tout cela ne contredit nullement le théorème de Maschke. En effet, si K est infini, le groupe G est infini. Si K est fini, G est fini d'ordre $|K|$, et dans ce cas, $\mathrm{car}(K) \mid |G|$.

5.4. Remarque. Notons que le théorème de Maschke devient faux si l'on remplace $K[G]$ par un anneau A quelconque, même pour un anneau aussi simple que $\mathbb{Z}$. En effet, des arguments semblables à ceux développés dans l'exemple 1.5 (2) montrent que le sous-module $2\mathbb{Z}$ de $\mathbb{Z}$ ne possède pas de supplémentaire.

Détaillons maintenant un exemple.

5.5. Exemple. Soit K un corps de caractéristique différente de 2 et 3. Considérons la représentation régulière de $\mathfrak{S}_3$. C'est simplement la représentation de permutation $K[\mathfrak{S}_3]$, associée à l'action de $\mathfrak{S}_3$ sur lui-même par translation à gauche. D'après le théorème de Maschke, on sait que $K[\mathfrak{S}_3]$ est une somme directe interne de sous-représentations irréductibles, et par conséquent isomorphe à une somme directe externe de représentations irréductibles du groupe $\mathfrak{S}_3$.

Pour l'instant, nous connaissons trois exemples de représentations irréductibles[2] de $\mathfrak{S}_3$: la représentation triviale, la représentation donnée par la signature et la représentation H_3. Nous allons donc chercher des sous-représentations de $K[\mathfrak{S}_3]$ isomorphes à chacune de ces représentations irréductibles (si elles existent !).

2. Nous verrons plus tard que ce sont en fait les seules.

Posons $c = (1\ 2\ 3)$ et $t = (2\ 3)$. Rappelons que c est d'ordre 3, t est d'ordre 2, et $tct^{-1} = c^2$. La famille $\mathbf{e} = (e_{\mathrm{Id}}, e_c, e_{c^2}, e_t, e_{ct}, e_{c^2t})$ est alors une base de $K[\mathfrak{S}_3]$.

Une sous-représentation de $K[\mathfrak{S}_3]$ est isomorphe à la représentation triviale si c'est un sous-espace de dimension 1 sur lequel $\mathfrak{S}_3$ agit trivialement. Or, $K[\mathfrak{S}_3]^{\mathfrak{S}_3}$ est un tel sous-espace d'après l'exemple 2.13, puisque l'action de $\mathfrak{S}_3$ sur lui-même par translation à gauche est transitive. On vérifie d'ailleurs que $K[\mathfrak{S}_3]^{\mathfrak{S}_3}$ est engendré par

$$f_1 = e_{\mathrm{Id}} + e_c + e_{c^2} + e_t + e_{ct} + e_{c^2t}.$$

Si l'on pose $V_1 = \mathrm{Vect}_K(f_1)$, on a donc

$$V_1 \simeq (K, \mathbf{1}),$$

par construction.

Une sous-représentation de $K[\mathfrak{S}_3]$ est isomorphe à la représentation donnée par la signature si c'est un sous-espace de dimension 1, engendré par un élément f_2 tel que

$$c\cdot f_2 = f_2 \quad \text{et} \quad t\cdot f_2 = -f_2.$$

Des calculs immédiats montrent que l'ensemble des fonctions vérifiant ces deux relations forment un sous-espace de dimension 1, engendré par

$$f_2 = e_{\mathrm{Id}} + e_c + e_{c^2} - e_t - e_{ct} - e_{c^2t}.$$

En posant $V_\varepsilon = \mathrm{Vect}_K(f_2)$, on a donc

$$V_\varepsilon \simeq (K, \varepsilon).$$

Enfin, en contemplant la représentation matricielle de H_3 décrite dans l'exemple 2.19 (4), on constate qu'une sous-représentation de $K[\mathfrak{S}_3]$ est isomorphe à H_3 si elle est de dimension 2, et engendrée par deux fonctions φ_1, φ_2 vérifiant les relations

$$\begin{aligned} c\cdot\varphi_1 &= -\varphi_2, & c\cdot\varphi_2 &= \varphi_1 - \varphi_2, \\ t\cdot\varphi_1 &= \varphi_2, & t\cdot\varphi_2 &= \varphi_1. \end{aligned}$$

Des calculs un peu longuets mais élémentaires montrent que de telles fonctions φ_1 et φ_2 sont de la forme

$$\begin{aligned} \varphi_1 &= x_1(e_{\mathrm{Id}} - e_{c^2t}) + x_2(e_c - e_{ct}) + x_3(e_{c^2} - e_t) \\ \varphi_2 &= x_1(e_t - e_c) + x_2(e_{c^2t} - e_{c^2}) + x_3(e_{ct} - e_{\mathrm{Id}}), \end{aligned}$$

avec $x_1, x_2, x_3 \in K$ vérifiant $x_1 + x_2 + x_3 = 0$.

On définit alors quatre éléments de $K[\mathfrak{S}_3]$ par

$$\begin{aligned} f_3 &= e_{\mathrm{Id}} - e_c + e_{ct} - e_{c^2t} \\ f_4 &= -e_c + e_{c^2} + e_t - e_{c^2t} \\ f_5 &= e_{\mathrm{Id}} - e_{c^2} + e_t - e_{c^2t} \\ f_6 &= e_{\mathrm{Id}} - e_c + e_t - e_{ct}. \end{aligned}$$

Si l'on pose $V_3 = \mathrm{Vect}_K(f_3, f_4)$, et $V_3' = \mathrm{Vect}_K(f_5, f_6)$, les représentations linéaires V_3 et V_3' sont alors isomorphes à la représentation H_3 par construction.

On vérifie alors sans mal que le déterminant de la famille $(f_1, \ldots, f_6)$ dans la base $\mathbf{e}$ est égal à -54, qui est non nul grâce à l'hypothèse faite sur la caractéristique de K. C'est donc une base de $K[\mathfrak{S}_3]$. Il s'ensuit que l'on a une décomposition en somme directe de sous-représentations irréductibles

$$K[\mathfrak{S}_3] = V_1 \oplus V_\varepsilon \oplus V_3 \oplus V_3'.$$

On a donc un isomorphisme de représentations

$$K[\mathfrak{S}_3] \simeq (K, \mathbf{1}) \times (K, \varepsilon) \times H_3 \times H_3.$$

Nous verrons plus loin une méthode plus systématique pour trouver la décomposition d'une représentation donnée.

6. Exercices

E1. Soit K un corps et soit $\varphi : G \longrightarrow H$ un morphisme de groupes.

a. Montrer que φ s'étend de manière unique en un morphisme d'algèbres

$$K[\varphi] : K[G] \longrightarrow K[H].$$

b. Montrer que $K[\varphi]$ est injectif (resp. surjectif, bijectif) si, et seulement si, φ l'est. En déduire que deux groupes isomorphes ont des algèbres de groupe isomorphes.

c. Soit $(g_i)_{i \in I}$ un système complet de représentants de $G/\mathrm{Ker}(\varphi)$. Montrer que les éléments

$$\sum_{h \in \mathrm{Ker}(\varphi) \setminus \{1_G\}} (e_{g_i h} - e_{g_i}), \ i \in I$$

forment une K-base de $\mathrm{Ker}(K[\varphi])$.

E2. Soit K un corps de caractéristique différente de 2, et contenant un élément $i \in K^\times$ tel que $i^2 = -1$. Soit G un groupe cyclique d'ordre 4, engendré par g.

On pose

$$\alpha = \frac{1+i}{2} e_g + \frac{1-i}{2} e_{g^3}, \quad \beta = \frac{1-i}{2} e_g + \frac{1+i}{2} e_{g^3}.$$

a. Montrer que $H = \{e_1, e_{g^2}, \alpha, \beta\}$ est un sous-groupe de $K[G]^\times$ isomorphe $(\mathbb{Z}/2\mathbb{Z})^2$.

b. Montrer que les K-algèbres $K[G]$ et $K[H]$ sont isomorphes.

E3. Soit $n \geqslant 2$. On suppose que $\mathrm{car}(K) \nmid n$, et l'on munit K^n de l'action de $\mathfrak{S}_n$ définie par

$$\begin{aligned} \mathfrak{S}_n \times K^n &\longrightarrow K^n \\ (\sigma, (x_i)_{i\in[\![1,n]\!]}) &\longmapsto (x_{\sigma^{-1}(i)})_{i\in[\![1,n]\!]}. \end{aligned}$$

On souhaite démontrer que les sous-représentations de K^n sont

$$0, \quad K^n, \quad H_n \quad \text{et} \quad K\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix},$$

où $H_n = \big\{(x_i)_{i\in[\![1,n]\!]} \in K^n \mid x_1 + \cdots + x_n = 0\big\}$.

Soit V une sous-représentation de K^n.

a. On suppose que $V \subset H_n$. Montrer que $V = \{0\}$ ou $V = H_n$.

b. On suppose maintenant que V n'est pas contenu dans H_n.

(i) En considérant l'action des puissances de $(1\ 2 \cdots\ n)$ sur un vecteur de V bien choisi, montrer que $v_0 = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix} \in V$.

(ii) Montrer pour tout $v \in V$, il existe $\lambda \in K$ tel que $v - \lambda v_0 \in H_n$, puis que $V = (V \cap H_n) \oplus Kv_0$.

(iii) En déduire que $V = Kv_0$ ou $V = K^n$.

E4. Soit G un groupe fini. Montrer qu'une représentation irréductible de G est de degré fini.

Indication. Soit V une représentation irréductible, et soit $v \in V \setminus \{0\}$. Que dire de $W = \mathrm{Vect}_K(g{\cdot}v,\ g \in G)$?

E5. Soit K un corps de caractéristique $p > 0$ et soit G un p-groupe fini.

a. Soit V une représentation non nulle de G (non nécessairement de dimension finie sur K).

Soit $v \in V \setminus \{0\}$. En faisant agir G sur $E = \mathrm{Vect}_{\mathbb{F}_p}(g{\cdot}v, g \in G) \setminus \{0\}$, en déduire que le sous-espace V^G des points fixes est non nul.

b. Quelles sont les représentations irréductibles de G ?

E6. Soit K un corps.

a. Montrer qu'il existe deux représentations $\rho_1, \rho_2 : \mathfrak{S}_3 \longrightarrow \mathrm{GL}_2(K)$ vérifiant

$$\rho_1\big((1\ 2)\big) = \begin{pmatrix} -1 & 1 \\ 0 & 1 \end{pmatrix}, \qquad \rho_1\big((2\ 3)\big) = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}$$

$$\rho_2\big((1\ 2)\big) = \begin{pmatrix} -2 & -1 \\ 3 & 2 \end{pmatrix}, \qquad \rho_2\big((2\ 3)\big) = \begin{pmatrix} 1 & 0 \\ -3 & -1 \end{pmatrix}.$$

b. Montrer que ρ_1 et ρ_2 sont isomorphes.

c. Montrer que ρ_1 est irréductible si, et seulement si, $\mathrm{car}(K) \neq 3$.

E7. Soit $n \geqslant 1$ un entier, soit K un corps tel que $\mathrm{car}(K) \nmid n$, et soit G un groupe abélien d'ordre n.

a. On suppose que $\mu_n(\overline{K}) \subset K$. Montrer que toute représentation de G de degré fini est somme directe interne de représentations de degré 1.

Indication. Si $\rho : G \longrightarrow \mathrm{GL}(V)$ est une représentation linéaire, montrer que l'endomorphisme $\rho(g)$ est diagonalisable pour tout $g \in G$, et utiliser l'exercice XXX-3.

b. En déduire que si $\mu_n(\overline{K}) \subset K$, les représentations irréductibles de G sont les représentations de degré 1. Le résultat est-il encore vrai sans cette hypothèse ?

E8. Soit V une représentation de G sur K et soit L/K une extension. Montrer que si $V \otimes_K L$ est irréductible, V est irréductible. La réciproque est-elle vraie ?

Indication. Raisonner par contraposition.

E9. Détailler l'exemple 4.13.

E10. Soit K un corps et soit G un groupe (pas nécessairement fini). On considère une représentation matricielle $\rho : G \longrightarrow \mathrm{GL}_n(K)$ non irréductible (en particulier, $n \geqslant 2$).

a. Montrer qu'il existe une base $\mathbf{e}$ de K^n et des entiers $n_1, n_2 \geqslant 1$ tels que, pour tout $g \in G$,

$$\mathrm{Mat}\big(\rho(g)\,;\mathbf{e}\big) = \begin{pmatrix} M_1(g) & N(g) \\ 0 & M_2(g) \end{pmatrix},$$

avec $M_1(g) \in \mathrm{GL}_{n_1}(K)$, $M_2(g) \in \mathrm{GL}_{n_2}(K)$ et $N(g) \in \mathrm{M}_{n_1 \times n_2}(K)$.

b. Pour $i \in \{1,2\}$, on pose $G_i = \{M_i(g) \mid g \in G\}$. Montrer que G_i est un sous-groupe de $\mathrm{GL}_{n_i}(K)$, et que l'application

$$\begin{aligned} \tau : G &\longrightarrow G_1 \times G_2 \\ g &\longmapsto \big(M_1(g), M_2(g)\big) \end{aligned}$$

est un morphisme de groupes[3].

c. Montrer que si G est fini et $\mathrm{car}(K) \nmid |G|$, on peut choisir $\mathbf{e}$ de telle sorte que $N(g) = 0$, pour tout $g \in G$. Est-ce encore le cas si G est infini ou si $\mathrm{car}(K) \mid |G|$?

E11. Soit K un corps. Pour tout $P \in K[X_1, \ldots, X_n]$, et tout $M \in \mathrm{GL}_n(K)$, on pose

$$M \bullet P = P\Big(\sum_{j=1}^{n} a_{1j}X_j, \ldots, \sum_{j=1}^{n} a_{nj}X_j\Big),$$

où $M^{-1} = (a_{ij})_{i,j}$.

On rappelle que cela définit une action de $\mathrm{GL}_n(K)$ sur $K[X_1, \ldots, X_n]$ par automorphismes d'algèbre (cf. exercice XXVI-13 pour une définition).

a. Pour tout $M \in \mathrm{GL}_n(K)$, et tout $P \in K[X_1, \ldots, X_n]$ homogène de degré $d \geqslant 0$, montrer que $M \bullet P$ est homogène de degré d.

Soit G un groupe fini, et soit $\rho : G \longrightarrow \mathrm{GL}_n(K)$ une représentation matricielle de G.

D'après la question précédente, l'action de G sur $K[X_1, \ldots, X_n]$ définie par

$$\begin{aligned} G \times K[X_1, \ldots, X_n] &\longrightarrow K[X_1, \ldots, X_n] \\ (g, P) &\longmapsto g{\cdot}P = \rho(g) \bullet P \end{aligned}$$

se restreint en une action sur le K-espace vectoriel $K[X_1, \ldots, X_n]_d$ des polynômes homogènes de degré d.

b. Si L/K est une extension de corps, pour tout $d \geqslant 0$, montrer que

$$\dim_L \big(L[X_1, \ldots, X_n]_d^G\big) = \dim_K \big(K[X_1, \ldots, X_n]_d^G\big),$$

où l'action de G sur $L[X_1, \ldots, X_n]$ est associée au même morphisme ρ, mais vu comme à valeurs dans $\mathrm{GL}_n(L)$.

Indication. Un polynôme P à coefficients dans K ou dans L est fixe sous l'action de G si, et seulement si, ses coefficients sont solutions d'un même système linéaire à coefficients dans K.

3. Ce fait sera utilisé dans certains des exercices qui suivent.

On se propose de démontrer le théorème de Molien : si K est de caractéristique 0, alors

$$\frac{1}{|G|}\sum_{g\in G}\frac{1}{\det\left(\mathrm{I}_n - T\rho(g)\right)} = \sum_{d\geqslant 0}\dim_K\left(K[X_1,\ldots,X_n]_d^G\right)T^d \in K[[T]].$$

c. Montrer que l'on peut supposer K algébriquement clos, ce qui l'on fera désormais.

d. Soit $g \in G$. Justifier que $\rho(g)$ est diagonalisable.

On écrit $\rho(g) = SDS^{-1}$, avec $D = \begin{pmatrix} \lambda_{1,g} & & \\ & \ddots & \\ & & \lambda_{n,g} \end{pmatrix}$.

e. Pour tout $i \in [\![1,n]\!]$, on pose $U_i = S \bullet X_i$. Montrer que les polynômes

$$S \bullet U_1^{\alpha_1}\cdots U_n^{\alpha_n}, \quad \alpha_1,\ldots,\alpha_n \in \mathbb{N},\ \alpha_1+\cdots+\alpha_n = d$$

forment une base de $K[X_1,\ldots,X_n]_d$.
Montrer que $\rho(g)^{-1} \bullet U_i = \lambda_{i,g}^{-1}U_i$, pour tout $i \in [\![1,n]\!]$.

f. Montrer que la trace de l'endomorphisme

$$\begin{array}{ccc} K[X_1,\ldots,X_n]_d & \longrightarrow & K[X_1,\ldots,X_n]_d \\ P & \longmapsto & g\cdot P \end{array}$$

est égale à $\displaystyle\sum_{\alpha_1+\cdots+\alpha_n=d}\lambda_{1,g}^{-\alpha_1}\cdots\lambda_{n,g}^{-\alpha_n}$.

g. En déduire que, pour tout $d \geqslant 0$, on a

$$\dim_K(K[X_1,\ldots,X_n]_d^G) = \frac{1}{|G|}\sum_{g\in G}\sum_{\alpha_1+\cdots+\alpha_n=d}\lambda_{1,g}^{-\alpha_1}\cdots\lambda_{n,g}^{-\alpha_n}.$$

h. Justifier que la fraction rationnelle $\dfrac{1}{\det\left(\mathrm{I}_n - T\rho(g)^{-1}\right)}$ est développable en série formelle, et que l'on a

$$\frac{1}{\det\left(\mathrm{I}_n - T\rho(g)^{-1}\right)} = \prod_{i=1}^{n}\frac{1}{(1-\lambda_{i,g}^{-1}T)}.$$

Développer en série formelle, et en déduire le théorème de Molien.

i. En appliquant le théorème de Molien au morphisme trivial, démontrer que l'on a

$$\dim_K(K[X_1,\ldots,X_n]_d) = \binom{n+d-1}{d}, \quad \text{pour tout } d \geqslant 0.$$

Retrouver cette égalité par un raisonnement combinatoire.

Les résultats des exercices qui suivent sont dûs à Burnside.

E12. Soit K un corps algébriquement clos, soit $n \geqslant 1$, et soit $\mathscr{A}$ une sous-algèbre de $\mathrm{M}_n(K)$. On fait agir les éléments de $\mathscr{A}$ de manière naturelle sur K^n. On suppose que K^n est un $\mathscr{A}$-module simple. Le but de cet exercice est de démontrer que $\mathscr{A} = \mathrm{M}_n(K)$.

Si $M \in \mathrm{M}_n(K)$, on note $C_j(M)$ la j-ème colonne de M. Pour tout entier $k \in [\![1, n]\!]$, on pose

$$\mathscr{C}_k = \{M \in \mathrm{M}_n(K) \mid C_j(M) = 0, \quad \text{pour tout } j \neq k\}.$$

Si $X \in K^n$, on note $L_k(X) \in \mathscr{C}_k$ la matrice dont toutes les colonnes sont nulles, sauf la k-ième colonne, qui vaut X.

a. Montrer que $\mathscr{C}_k$ est un $\mathscr{A}$-module, que l'application

$$L_k : K^n \longrightarrow \mathscr{C}_k$$

est un isomorphisme de $\mathscr{A}$-modules, et que $\mathscr{C}_k$ est simple.

b. Vérifier que $\mathrm{M}_n(K) = \mathscr{C}_1 \oplus \cdots \oplus \mathscr{C}_n$.

c. Soit W un sous-module de $\mathrm{M}_n(K)$. Le but de cette question est de démontrer que W admet un $\mathscr{A}$-module supplémentaire.

(i) On suppose qu'il existe $k \in [\![1, n]\!]$ tel que $W \cap \mathscr{C}_k \neq 0$. Montrer que $W = \mathscr{C}_k$ et en déduire que W possède un supplémentaire.

(ii) On suppose qu'il existe $k \in [\![1, n]\!]$ tel que $W \cap \mathscr{C}_k = 0$. Justifier l'existence d'un sous-module W' non nul de dimension maximale tel que W et W' soient en somme directe. Montrer que l'on a

$$(W \oplus W') \cap \mathscr{C}_k \neq 0, \quad \text{pour tout } k \in [\![1, n]\!].$$

En déduire $\mathscr{C}_k \subset W \oplus W'$ pour tout $k \in [\![1, n]\!]$, puis que

$$\mathrm{M}_n(K) = W \oplus W'.$$

(iii) Conclure.

Pour tout $k \in [\![1, n]\!]$, et pour tout $M \in \mathrm{M}_n(K)$, on note $P_k(M)$ la matrice dont toutes les colonnes sont nulles, sauf la k-ième colonne, qui vaut $C_k(M)$.

d. Montrer que $P_k : \mathrm{M}_n(K) \longrightarrow \mathscr{C}_k$ est $\mathscr{A}$-linéaire.

e. Soit W' un $\mathscr{A}$-module supplémentaire de $\mathscr{A}$ dans $\mathrm{M}_n(K)$, et soit

$$\pi : \mathrm{M}_n(K) \longrightarrow \mathrm{M}_n(K)$$

le projecteur sur $\mathscr{A}$ parallèlement à W'. On note $(\varepsilon_1, \ldots, \varepsilon_n)$ la base canonique de K^n.

Vérifier que $\pi(\mathrm{I}_n) = \mathrm{I}_n$, que $\sum_{k=1}^{n} L_k(\varepsilon_k) = \mathrm{I}_n$ et que $P_k(\mathrm{I}_n) = L_k(\varepsilon_k)$, pour tout $k \in [\![1, n]\!]$.

f. En utilisant le lemme de Schur, montrer que pour tous i, $j \in [\![1, n]\!]$, il existe $a_{ij} \in K$ tel que

$$L_i^{-1} \circ P_j \circ \pi \circ L_i = a_{ij} \mathrm{Id}_{K^n}.$$

g. Montrer alors que $a_{ij} = \delta_{ij}$, pour tous i, $j \in [\![1, n]\!]$.

Indication. On pourra considérer $\sum_{i=1}^{n} a_{ij} \varepsilon_i$.

h. Montrer que $\pi \circ L_k = L_k$, puis que $\mathscr{C} \subset \mathscr{A}$, pour tout $k \in [\![1, n]\!]$. Conclure.

Indication. On pourra remarquer que $\sum_{k=1}^{n} P_k = \mathrm{Id}_{\mathrm{M}_n(k)}$ et utiliser les questions précédentes.

E13. Soit K un corps algébriquement clos et soit G un groupe (pas nécessairement fini). On se propose de démontrer qu'une représentation matricielle $\rho : G \longrightarrow \mathrm{GL}_n(K)$ est irréductible si, et seulement si, les éléments de $\rho(G)$ forment une famille génératrice de $\mathrm{M}_n(K)$.

a. On suppose que K^n est une représentation irréductible de G.

(i) Justifier que K^n est une représentation irréductible de $\rho(G)$.

(ii) En utilisant l'exercice précédent avec une sous-algèbre de $\mathrm{M}_n(K)$ bien choisie, montrer que les éléments de $\rho(G)$ forment une famille génératrice de $\mathrm{M}_n(K)$.

b. On suppose que K^n n'est pas irréductible. Utiliser l'exercice 10 pour montrer que les éléments de $\rho(G)$ n'engendrent pas $\mathrm{M}_n(K)$.

c. Conclure.

E14. Soit K un corps algébriquement clos et soit $n \geqslant 1$. Soit G un sous-groupe de $\mathrm{GL}_n(K)$. On suppose que K^n est une représentation irréductible de G (pour l'action naturelle de G sur K^n).

a. Soit $(g_1, \ldots, g_{n^2})$ une base de $\mathrm{M}_n(K)$ formée d'éléments de G (une telle base existe d'après l'exercice précédent). Montrer que l'application

$$\begin{aligned} \tau : \mathrm{M}_n(K) &\longrightarrow K^{n^2} \\ M &\longmapsto (\mathrm{tr}(Mg_1), \ldots, \mathrm{tr}(Mg_{n^2})) \end{aligned}$$

est injective.

b. On suppose que l'application $\mathrm{tr} : G \longrightarrow K$ a une image finie, de cardinal r. Montrer que $|G| \leqslant r^{n^2}$.

E15. Soit $n \geqslant 1$, soit G un sous-groupe de $\mathrm{GL}_n(K)$ d'exposant fini e. On suppose que $\mathrm{car}(K) \nmid e$. Le but de cet exercice est de démontrer que l'on a $|G| \leqslant e^{n^3}$. En particulier, G est fini.

a. Justifier que l'on peut supposer que K est algébriquement clos.

On suppose que K est algébriquement clos, et l'on procède par récurrence sur n.

b. Traiter le cas $n = 1$.

On se fixe un entier $n \geqslant 2$, et l'on suppose que le résultat est vrai pour tout $m \in [\![1, n-1]\!]$, et pour tout sous-groupe de $\mathrm{GL}_m(K)$ d'exposant fini premier à $\mathrm{car}(K)$.

On considère alors un sous-groupe G de $\mathrm{GL}(K)$ d'exposant fini e tel que $\mathrm{car}(K) \nmid e$.

c. Montrer que pour tout $g \in G$, $\mathrm{tr}(g)$ est la somme de n racines e-ièmes de l'unité. En déduire que l'image de $\mathrm{tr} : G \longrightarrow K$ possède au plus e^n éléments.

d. On suppose que K^n est une représentation irréductible de G (pour l'action naturelle de G sur K^n). Utiliser l'exercice 14 pour montrer que $|G| \leqslant e^{n^3}$.

On suppose maintenant que K^n n'est pas irréductible.

D'après l'exercice 10, il existe alors une base $\mathbf{e}$ de l'espace K^n et deux entiers n_1, $n_2 \geqslant 1$ tels que, pour tout $g \in G$,

$$\mathrm{Mat}(g\,;\mathbf{e}) \;=\; \begin{pmatrix} M_1(g) & N(g) \\ 0 & M_2(g) \end{pmatrix},$$

avec $M_1(g) \in \mathrm{GL}_{n_1}(K)$, $M_2(g) \in \mathrm{GL}_{n_2}(K)$ et $N(g) \in \mathrm{M}_{n_1 \times n_2}(K)$.
On rappelle que $G_i = \{M_i(g) \mid g \in G\}$ est un sous-groupe de $\mathrm{GL}_{n_i}(K)$, et que l'application

$$\begin{aligned} \tau : G &\longrightarrow G_1 \times G_2 \\ g &\longmapsto \big(M_1(g), M_2(g)\big) \end{aligned}$$

est un morphisme de groupes.

e. Montrer que $\tau : G \longrightarrow G_1 \times G_2$ est injectif.

Indication. Si $g \in \mathrm{Ker}(\tau)$, calculer g^e et en déduire que $N(g) = 0$.

f. Achever la récurrence, et conclure.

g. Montrer que le résultat devient faux si $\mathrm{car}(K) \mid e$.

Indication. Penser à l'exemple 5.3.

E16. Soit $n \geqslant 1$ un entier, et soit G un sous-groupe de $\mathrm{GL}_n(K)$. Le but de cet exercice est de démontrer que G est fini si, et seulement si, G possède un nombre fini de classes de conjugaison [4] dans G.

a. Justifier que l'on peut supposer que K est algébriquement clos.

On suppose que K est algébriquement clos, et l'on procède par récurrence sur n.

b. Traiter le cas $n = 1$.

On se fixe un entier $n \geqslant 2$, et l'on suppose que le résultat est vrai pour tout $m \in [\![1, n-1]\!]$, et pour tout sous-groupe de $\mathrm{GL}_m(K)$ ayant un nombre fini de classes de conjugaison.
On considère alors un sous-groupe G de $\mathrm{GL}_n(K)$ ayant un nombre fini de classes de conjugaison.

c. Montrer que l'image de $\mathrm{tr} : G \longrightarrow K$ est finie.

d. On suppose que K^n est une représentation irréductible de G (pour l'action naturelle de G sur K^n). Utiliser l'exercice 14 pour montrer que G est un groupe fini.

On suppose maintenant que K^n n'est pas irréductible.
D'après l'exercice 10, il existe alors une base l'espace $\mathbf{e}$ de K^n et deux entiers n_1, $n_2 \geqslant 1$ tels que, pour tout $g \in G$,

$$\mathrm{Mat}(g\,;\mathbf{e}) \;=\; \begin{pmatrix} M_1(g) & N(g) \\ 0 & M_2(g) \end{pmatrix},$$

avec $M_1(g) \in \mathrm{GL}_{n_1}(K)$, $M_2(g) \in \mathrm{GL}_{n_2}(K)$ et $N(g) \in \mathrm{M}_{n_1 \times n_2}(K)$.
On rappelle que $G_i = \{M_i(g) \mid g \in G\}$ est un sous-groupe de $\mathrm{GL}_{n_i}(K)$, et que l'application

$$\begin{aligned} \tau : G &\longrightarrow G_1 \times G_2 \\ g &\longmapsto \big(M_1(g), M_2(g)\big) \end{aligned}$$

est un morphisme de groupes.

e. Montrer que G_i est un sous-groupe de $\mathrm{GL}_{n_i}(K)$ ayant un nombre fini de classes de conjugaison dans G_i.

f. Montrer que $\tau(G)$ est fini, et que $H = \mathrm{Ker}(\tau)$ est un groupe abélien d'indice fini dans G.

g. En déduire que tout élément $h \in H$ possède un nombre fini de conjugués dans G. Utiliser l'hypothèse sur G pour en déduire que H est fini. Montrer alors que G est fini, et conclure.

---

4. L'hypothèse que G est un sous-groupe de $\mathrm{GL}_n(K)$ est cruciale. On peut en effet exhiber des exemples de groupes infinis ne possédant qu'un nombre fini de classes de conjugaison. Néanmoins, la construction de ces exemples est très élaborée, et dépasse le cadre de cet ouvrage.

Chapitre XXXV

Théorie des caractères

Dans ce chapitre, on introduit la notion de caractère d'une représentation linéaire de degré fini, qui est la notion centrale de la théorie. On reprend la notation 1_K pour désigner l'élément unité du corps K.

1. Caractère d'une représentation

Dans tout ce qui suit, G est un groupe, K est un corps, et toute représentation V de G est de degré fini.

Introduisons sans attendre l'objet de notre étude.

1.1. Définition. Le *caractère* d'une représentation $\rho : G \longrightarrow \mathrm{GL}(V)$ de degré fini est l'application

$$\begin{aligned} G &\longrightarrow K \\ g &\longmapsto \mathrm{tr}\big(\rho(g)\big). \end{aligned}$$

On dit qu'une fonction $\chi \in \mathscr{F}(G, K)$ est un *caractère* (resp. un *caractère irréductible*) de G si c'est le caractère d'une représentation (resp. d'une représentation irréductible) de degré fini du groupe G.

On notera $\mathrm{Irr}(G)$ l'ensemble des caractères irréductibles du groupe G.

1.2. Remarques

(1) Si χ est le caractère d'une représentation (V, ρ) de degré fini, on a

$$\chi(1_G) = \mathrm{tr}\big(\rho(1_G)\big) = \mathrm{tr}(\mathrm{Id}_V) = \dim_K(V) \cdot 1_K.$$

(2) Deux représentations isomorphes de G de degré fini ont des caractères égaux, puisque deux endomorphismes semblables ont même trace.

L'exemple des représentations

$$\begin{array}{ccc} \mathbb{Z}/2\mathbb{Z} & \longrightarrow & \mathrm{GL}_2(K) \\ \overline{m} & \longmapsto & \mathrm{Id}_{K^2} \end{array} \qquad \text{et} \qquad \begin{array}{ccc} \mathbb{Z}/2\mathbb{Z} & \longrightarrow & \mathrm{GL}_2(K) \\ \overline{m} & \longmapsto & \begin{pmatrix} \overline{1} & \overline{m} \\ \overline{0} & \overline{1} \end{pmatrix}, \end{array}$$

où $\mathrm{car}(K) = 2$, montre que la réciproque est fausse en général, puisque celles-ci ont des caractères tous deux identiquement nuls, mais ne sont pas isomorphes (la première correspond à une action triviale sur K^2, contrairement à la seconde).

Nous verrons néanmoins plus loin que la réciproque devient vraie sous des hypothèses raisonnables.

(3) Soit $\chi : G \longrightarrow K$ un caractère de G. Les propriétés de la trace montrent que l'on a
$$\chi(gg'g^{-1}) = \chi(g'), \quad \text{pour tous } g,\, g' \in G.$$

Un caractère est donc constant sur les classes de conjugaison de G.

1.3. Définition. Une fonction $f \in \mathscr{F}(G, K)$ vérifiant
$$f(gg'g^{-1}) = f(g'), \quad \text{pour tous } g,\, g' \in G$$
est appelée une *fonction centrale sur* G de G.

On note $\mathscr{F}_c(G, K)$ l'ensemble des fonctions centrales sur G. C'est clairement un sous-espace vectoriel de $\mathscr{F}(G, K)$.

Le lemme suivant explique le nom de « fonction centrale ».

1.4. Lemme. *Soit G un groupe et soit K un corps. Alors, le centre de $K[G]$ est égal $\mathscr{F}_c(G, K)$.*

De plus, si G est fini, et si $\mathrm{Conj}(G)$ est l'ensemble des classes de conjugaison de G, pour tout $C \in \mathrm{Conj}(G)$, la fonction
$$C \longmapsto e_C = \sum_{g \in C} e_g$$
est une fonction centrale, et la famille $(e_C)_{C \in \mathrm{Conj}(G)}$ est alors une K-base de $Z(K[G])$.

En particulier, on a
$$\dim_K \left(Z(K[G])\right) = \dim_K \left(\mathscr{F}_c(G, K)\right) = |\mathrm{Conj}(G)|.$$

Démonstration. Soit $f = \sum_{h \in G} \lambda_h e_h \in K[G]$. Alors, $f \in Z(K[G])$ si, et seulement si, f commute avec e_g pour tout $g \in G$. Comme e_g est inversible, d'inverse $e_{g^{-1}}$, cela revient à dire que $e_{g^{-1}} f e_g = f$ pour tout $g \in G$.

Or, on a $$e_{g^{-1}} f e_g = \sum_{h \in G} \lambda_h e_{g^{-1}hg} = \sum_{h \in G} \lambda_{ghg^{-1}} e_h.$$

On a donc $e_g f = f e_g$ pour tout $g \in G$ si, et seulement si,

$$\lambda_{ghg^{-1}} = \lambda_h, \quad \text{pour tous } g, h \in G.$$

Comme $\lambda_h = f(h)$ pour tout $h \in G$, cela équivaut à dire que f est une fonction centrale.

Supposons maintenant que G soit fini. En particulier, toute classe de conjugaison est finie, et $\mathrm{Conj}(C)$ est fini. L'égalité précédente se réinterprète alors comme suit : pour tout $C \in \mathrm{Conj}(G)$, et pour tous h, $h' \in C$, on a $\lambda_h = \lambda_{h'}$. Notons λ_C cette valeur commune. On obtient alors que $f \in Z(K[G])$ si, et seulement si, $f = \sum_{C \in \mathrm{Conj}(G)} \lambda_C e_C$, pour certains $\lambda_C \in K$.

Ainsi, $(e_C)_{C \in \mathrm{Conj}(G)}$ est une famille génératrice de $Z(K[G])$. Puisque les classes de conjugaison forment une partition de G, les divers éléments e_C contiennent des vecteurs de base distincts, et l'on en déduit aisément que cette famille est libre. Cela achève la démonstration. □

Passons maintenant aux exemples.

1.5. Exemples

(1) Le caractère d'une représentation $\rho : G \longrightarrow K^\times$ de degré 1 est simplement l'application

$$\begin{aligned} G &\longrightarrow K \\ g &\longmapsto \rho(g). \end{aligned}$$

Attention néanmoins, cette application n'est plus un morphisme de groupes !

(2) Le caractère de la représentation triviale est noté χ_0. C'est simplement la fonction

$$\begin{aligned} \chi_0 : G &\longrightarrow K \\ g &\longmapsto 1_K. \end{aligned}$$

Plus généralement, si G agit trivialement sur V, le caractère associé est la fonction $\dim_K(V)\chi_0$.

(3) Si χ est le caractère d'une représentation V de G, alors pour tout caractère linéaire $\varphi : G \longrightarrow K^\times$, il est immédiat que le caractère de $V(\varphi)$ est $\varphi\chi$.

(4) Soit E un ensemble fini sur lequel G agit. Le caractère χ_E de la représentation de permutation $K[E]$ est, d'après l'exemple XXXIV-2.13,

$$\begin{aligned} \chi_E : G &\longrightarrow K \\ g &\longmapsto |\mathrm{Fix}(g)|. \end{aligned}$$

(5) Soit G un groupe fini agissant sur lui-même par translation à gauche. Tout élément $g \in G \setminus \{1_G\}$ agit sans point fixe, tandis que 1_G possède $|G|$ points fixes. Le caractère χ_{reg} de la représentation régulière est donc donné par

$$\chi_{\text{reg}}(g) = \begin{cases} |G| & \text{si } g = 1_G \\ 0 & \text{sinon.} \end{cases}$$

(6) Soit V une représentation de degré fini, de caractère χ. Le caractère χ^* de la représentation duale V^* est l'application

$$\begin{aligned} \chi^* : G &\longrightarrow K \\ g &\longmapsto \chi(g^{-1}). \end{aligned}$$

En effet, si $\mathbf{e}$ est une base de V, alors

$$\mathrm{Mat}\big(\rho(g^{-1})^*; \mathbf{e}^*\big) = \mathrm{Mat}\big(\rho(g^{-1}); \mathbf{e}\big)^t.$$

(7) Soient $\rho_1 : G \longrightarrow \mathrm{GL}(V_1)$ et $\rho_2 : G \longrightarrow \mathrm{GL}(V_2)$ deux représentations de degré fini, de caractères χ_{ρ_1} et χ_{ρ_2} respectivement. Alors, on a

$$\chi_{\rho_1 \oplus \rho_2} = \chi_{\rho_1} + \chi_{\rho_2} \quad \text{et} \quad \chi_{\rho_1 \otimes \rho_2} = \chi_{\rho_1} \chi_{\rho_2}.$$

Le résultat est clair d'après la description des matrices représentatives des endomorphismes $(\rho_1 \oplus \rho_2)(g)$ et $(\rho_1 \otimes \rho_2)(g)$.

(8) On suppose que $\mathrm{car}(K) \neq 2$. Soit V une représentation de degré fini, et soient χ_s et χ_a les caractères de $\mathrm{Sym}^2(V)$ et $\mathrm{Alt}^2(V)$. Alors, on a

$$\chi_s(g) = \frac{\chi(g)^2 + \chi(g^2)}{2} \quad \text{et} \quad \chi_a(g) = \frac{\chi(g)^2 - \chi(g^2)}{2},$$

pour tout $g \in G$.

Calculons χ_s. Soit $g \in G$ et soit $\mathbf{e} = (e_1, \ldots, e_n)$ une base de V. Alors, la famille $(e_i \otimes e_j + e_j \otimes e_i)_{i \leqslant j}$ est une base de $\mathrm{Sym}^2(V)$. Soit $g \in G$. Si $\rho(g)(e_\ell) = \sum_{k=1}^{n} a_{k\ell} e_k$, on a

$$\big(\rho(g) \otimes \rho(g)\big)(e_i \otimes e_j + e_j \otimes e_i) = \sum_{k,\ell} (a_{ki} a_{\ell j} + a_{\ell i} a_{kj})(e_k \otimes e_\ell + e_\ell \otimes e_k),$$

et ainsi

$$\begin{aligned} \chi_s(g) &= \sum_{i \leqslant j} (a_{ii} a_{jj} + a_{ji} a_{ij}) \\ &= \frac{1}{2} \sum_{i,j} (a_{ii} a_{jj} + a_{ji} a_{ij}) \\ &= \frac{1}{2} \Big(\Big(\sum_i a_{ii} \Big)^2 + \sum_{i,j} a_{ij} a_{ji} \Big) \\ &= \frac{1}{2} \big(\mathrm{tr}\big(\rho(g)\big)^2 + \mathrm{tr}\big(\rho(g)^2\big) \big) \\ &= \frac{1}{2} \big(\mathrm{tr}\big(\rho(g)\big)^2 + \mathrm{tr}\big(\rho(g^2)\big) \big) \\ &= \frac{1}{2} \big(\chi(g)^2 + \chi(g^2) \big). \end{aligned}$$

Comme $V \otimes_K V = \mathrm{Sym}^2(V) \oplus \mathrm{Alt}^2(V)$, le point (7) montre alors l'égalité

$$\chi^2 = \chi_s + \chi_a,$$

d'où la formule souhaitée pour χ_a.

(9) On suppose que G est fini et que $\mathrm{car}(K) \nmid |G|$. Soit V^G le sous-espace des points fixes et soit

$$V_G = \Big\{ v \in V \;\Big|\; \sum_{g \in G} g \cdot v = 0 \Big\}.$$

D'après l'exemple XXXIV-4.3 (2), on a $V = V^G \oplus V_G$. Puisque G agit trivialement sur V^G, le caractère de V^G est donc $\dim_K(V^G)\chi_0$. Si χ est le caractère de V et χ' celui de V_G, on a donc

$$\chi = \dim_K(V^G)\chi_0 + \chi'.$$

Autrement dit, on a

$$\chi'(g) = \chi(g) - \dim_K(V^G) \cdot 1_K, \quad \text{pour tout } g \in G.$$

(10) Si χ est le caractère d'une représentation V de G, pour tout sous-groupe H de G, le caractère de $\mathrm{Res}^G_H(V)$ est $\chi_{|H}$. Cela découle des définitions.

(11) On suppose que G est fini et que $\mathrm{car}(K) \nmid |G|$. Soit E un ensemble fini non vide sur lequel G agit. Notons χ'_E le caractère de $K[E]_G$, avec les notations précédentes. L'exemple XXXIV-2.13 et le calcul du caractère χ_E de $K[E]$ montrent alors que

$$\chi'_E(g) = \big(|\mathrm{Fix}(g)| - |\Omega|\big) \cdot 1_K, \quad \text{pour tout } g \in G,$$

où Ω est l'ensemble des G-orbites.

En particulier, pour tout $n \geqslant 2$, le caractère χ_{H_n} associé à H_n est donné par

$$\chi_{H_n}(\sigma) = \big(|\mathrm{Fix}(\sigma)| - 1\big) \cdot 1_K,$$

l'action de $\mathfrak{S}_n$ sur $E = [\![1, n]\!]$ étant transitive. Nous avons ici utilisé le fait que H_n était isomorphe à $K[E]_{\mathfrak{S}_n}$, et donc que les caractères de ces représentations sont les mêmes par la remarque 1.2 (1).

(12) Soit $\pi : G \longrightarrow G'$ un morphisme surjectif. Si χ' est le caractère d'une représentation V' de G', le caractère de l'inflation $\pi^*(V')$ est simplement $\chi' \circ \pi$.

Si V est une représentation de G sur laquelle $\mathrm{Ker}(\pi)$ agit trivialement, le caractère $\overline{\chi}$ de $\overline{V}$ est donné par

$$\overline{\chi}(\overline{g}) = \chi(g), \quad \text{pour tout } \overline{g} \in G/\mathrm{Ker}(\pi).$$

Là encore, cela découle directement des définitions.

(13) Soit L/K une extension de corps. Si V est une représentation de G de degré fini, alors V et $V \otimes_K L$ ont même caractère.
En effet, si $\rho : G \longrightarrow \mathrm{GL}(V)$ est le morphisme correspondant, et si $\mathbf{e}$ est une K-base de V, alors pour tout $g \in G$, on a $\rho_L(g) = \rho(g)_L$ et ainsi

$$\mathrm{Mat}\big(\rho_L(g); \mathbf{e} \otimes \mathbf{1}\big) \;=\; \mathrm{Mat}\big(\rho(g)\,; \mathbf{e}\big).$$

En particulier, $\mathrm{tr}\big(\rho_L(g)\big) = \mathrm{tr}\big(\rho(g)\big)$ pour tout $g \in G$, d'où l'assertion.

1.6. Remarque. Soit G un groupe et soit K un corps de caractéristique nulle. Soit (V, ρ) une représentation de G, de caractère χ. Alors, $\chi = \chi_0$ si, et seulement si, V est la représentation triviale.

En effet, pour tout $g \in G$, on a

$$\chi(1_G) \;=\; \dim_K(V) \cdot 1_K \;=\; \chi_0(1_G) \;=\; 1_K.$$

Le corps K étant de caractéristique nulle, on en déduit que $\dim_K(V) = 1$. Ainsi, après identification canonique de $\mathrm{GL}(V)$ avec $K^\times$, on peut considérer que ρ prend ses valeurs dans $K^\times$. Mais alors, pour tout $g \in G$, on a

$$\rho(g) \;=\; \chi(g) \;=\; \chi_0(g) = 1_K,$$

et V est la représentation triviale.

La détermination des caractères d'un groupe G (en particulier des caractères irréductibles) est une question vraiment cruciale. En effet, on verra que les représentations de G sont entièrement déterminées à isomorphisme près par leurs caractères lorsque K est de caractéristique nulle (cf. corollaire 2.12). Les propriétés des caractères permettent aussi de tirer quelques conséquences assez frappantes sur la structure de G. Afin d'illustrer notre propos, nous allons démontrer un critère de simplicité d'un groupe fini en termes de caractères irréductibles.

On commence par un lemme.

1.7. Lemme. *On suppose que $K = \mathbb{C}$. Soit G un groupe fini.*

Soit $\rho : G \longrightarrow \mathrm{GL}(V)$ une représentation de G, de caractère χ. Alors :

(1) *pour tout $g \in G$, $\rho(g)$ est diagonalisable et ses valeurs propres sont des racines d-ièmes de l'unité, où $d = o(g)$. En particulier, $\chi(g)$ est une somme de racines d-ièmes de l'unité (et donc de racines e-ièmes de l'unité), et $\chi(g)$ est entier sur $\mathbb{Z}$;*

(2) *soit $e = \exp(G)$. Pour tout $g \in G$, on a*

$$|\chi(g)| \leqslant \chi(1_G).$$

De plus, on a égalité si, et seulement si, il existe une racine e-ième de l'unité $z \in \mu_e(\mathbb{C})$ telle que $\rho(g) = z\mathrm{Id}_V$;

(3) *pour tout* $g \in G$, *on a* $\chi(g) = \chi(1_G)$ *si, et seulement si,* $\rho(g) = \mathrm{Id}_V$. *Autrement dit,*

$$\mathrm{Ker}(\rho) \;=\; \{g \in G \mid \chi(g) \;=\; \chi(1_G)\}.$$

En particulier :

(a) *le morphisme* $\rho : G \longrightarrow \mathrm{GL}(V)$ *est trivial si, et seulement si, pour tout* $g \in G$, *on a* $\chi(g) = \chi(1_G)$;

(b) *le morphisme* $\rho : G \longrightarrow \mathrm{GL}(V)$ *est fidèle si, et seulement si, on a*

$$\chi(g) \neq \chi(1_G), \qquad \textit{pour tout } g \in G \setminus \{1_G\}.$$

Démonstration. Soit $n = \dim_K(V) = \chi(1_G)$ et soit $g \in G$. Si $d = o(g)$, on a $g^d = 1$, et par conséquent

$$\rho(g)^d \;=\; \rho(g^d) \;=\; \rho(1_G) \;=\; \mathrm{Id}_V.$$

Ainsi, $X^d - 1$ annule $\rho(g)$. En particulier, $\rho(g)$ est diagonalisable et ses valeurs propres $z_1, \ldots, z_n \in \mathbb{C}$ (non nécessairement distinctes) sont des racines d-ièmes de l'unité. Mais alors, $\chi(g) = z_1 + \cdots + z_n$ est une somme de racines d-ièmes de l'unité, donc une somme d'éléments entiers sur $\mathbb{Z}$. Par conséquent, $\chi(g)$ est aussi entier sur $\mathbb{Z}$, par le corollaire XXVI-4.6. Notons que, puisque e est le ppcm des ordres des éléments de G, on a $d \mid e$, et toute racine d-ième de l'unité est aussi une racine e-ième de l'unité, d'où (1).

Montrons (2). L'inégalité triangulaire donne

$$|\chi(g)| \;=\; |z_1 + \cdots + z_n| \;\leqslant\; n,$$

puisque chaque z_i est de module 1. Supposons maintenant que $|\chi(g)| = n$. Autrement dit, $\dfrac{z_1 + \cdots + z_n}{n}$ est de module 1.

Pour tout entier $k \in [\![1, n]\!]$, écrivons $z_k = e^{i\theta_k}$, avec $\theta_k \in [0, 2\pi[$. De même, on écrit $\dfrac{z_1 + \cdots + z_n}{n} = e^{i\theta}$, avec $\theta \in [0, 2\pi[$.

On a donc $e^{i(\theta_1 - \theta)} + \cdots + e^{i(\theta_n - \theta)} = n$. En prenant la partie réelle, on obtient

$$\cos(\theta_1 - \theta) + \cdots + \cos(\theta_n - \theta) \;=\; n.$$

S'il existait $k \in [\![1, n]\!]$ tel que $\cos(\theta_k - \theta) < 1$, on aurait

$$\cos(\theta_1 - \theta) + \cdots + \cos(\theta_n - \theta) \;<\; n.$$

Par conséquent, pour tout $k \in [\![1, n]\!]$, on a $\cos(\theta_k - \theta) = 1$.

Comme $\theta_k - \theta \in]-2\pi, 2\pi[$, on en déduit que $\theta_k = \theta$, pour tout $k \in [\![1, n]\!]$. Ainsi, on a $z_k = z_1$, pour tout $k \in [\![1, n]\!]$. Comme $\rho(g)$ est diagonalisable, on obtient $\rho(g) = z_1 \mathrm{Id}_V$. Notons que $z_1 \in \mu_d(\mathbb{C}) \subset \mu_e(\mathbb{C})$.

Inversement, si on a $\rho(g) = z\mathrm{Id}_V$, avec $z \in \mu_e(\mathbb{C})$, on a

$$|\chi(g)| \;=\; |nz| \;=\; n.$$

Enfin, si $g \in \mathrm{Ker}(\rho)$, on a

$$\chi(g) \;=\; \mathrm{tr}\big(\rho(g)\big) \;=\; \mathrm{tr}(\mathrm{Id}_V) \;=\; \mathrm{tr}\big(\rho(1_G)\big) \;=\; \chi(1_G).$$

Inversement, soit $g \in G$ tel que $\chi(g) = \chi(1_G)$. Le point précédent montre alors que $\rho(g) = z\mathrm{Id}_V$, où $z \in \mu_e(\mathbb{C})$. On a ainsi

$$\chi(1_G) \;=\; \chi(g) \;=\; nz \;=\; \chi(1_G)z,$$

d'où $z = 1$ et $\rho(g) = \mathrm{Id}_V$.

La dernière partie du lemme est alors claire. Cela achève la démonstration.□

On a alors le théorème suivant.

1.8. Théorème. *On suppose que $K = \mathbb{C}$. Soit G un groupe fini non trivial. Alors, le groupe G est simple si, et seulement si, pour tout caractère irréductible $\chi \neq \chi_0$, et pour tout $g \in G \setminus \{1_G\}$, on a $\chi(g) \neq \chi(1_G)$ (c'est-à-dire si, et seulement si, toute représentation irréductible non triviale de G est fidèle).*

Démonstration. Soit G un groupe simple, et soit χ un caractère irréductible de G, associé à une représentation irréductible $\rho : G \longrightarrow \mathrm{GL}(V)$. Supposons qu'il existe $g \in G \setminus \{1_G\}$ tel que $\chi(g) = \chi(1_G)$. Le lemme 1.7 implique alors que $\mathrm{Ker}(\rho) \neq \{1_G\}$. Comme le groupe G est simple, on en déduit que $\mathrm{Ker}(\rho) = G$, et donc que ρ est le morphisme trivial. Autrement dit, on a $V \simeq \dim_K(V)\mathbf{1}$. Puisque V est une représentation irréductible, on en déduit que $V \simeq \mathbf{1}$, et donc que $\chi = \chi_0$.

Supposons maintenant que G ne soit pas simple. Nous allons construire un caractère irréductible $\chi \neq \chi_0$ tel qu'il existe un élément $g \in G \setminus \{1_G\}$ vérifiant $\chi(g) = \chi(1_G)$.

Puisque G n'est pas simple, il existe un sous-groupe H distingué dans G, distinct de $\{1_G\}$ et G. Le groupe G/H étant en particulier non trivial, il admet une représentation irréductible V correspondant à un morphisme de groupes $\rho' : G/H \longrightarrow \mathrm{GL}(V)$ non trivial d'après le corollaire XXXIV-5.2.

Soit $\pi : G \longrightarrow G/H$ la projection canonique. Alors, la représentation $\pi^*(V)$ est irréductible d'après l'exemple XXXIV-4.13, de caractère $\chi = \chi' \circ \pi$ d'après l'exemple 1.5(12). De plus, on a $\chi \neq \chi_0$. En effet, dans le cas contraire, χ' serait aussi le caractère associé à la représentation triviale de G/H, et ρ' serait en particulier triviale d'après la remarque 1.6, ce qui

n'est pas le cas. Choisissons maintenant $g \in H \setminus \{1_G\}$. On a alors

$$\chi(g) = \chi'(\overline{g}) = \chi'(\overline{1}_G) = \chi(1_G).$$

Cela achève la démonstration. □

Afin de motiver encore un peu plus l'étude des représentations des groupes finis et de leurs caractères, nous allons généraliser les considérations précédentes, et expliquer comment trouver tous les sous-groupes distingués d'un groupe fini à partir de ses caractères irréductibles.

Commençons par une définition.

1.9. Définition. Soit G un groupe, et soit χ un caractère de G. Le *noyau* de χ est l'ensemble

$$\mathrm{Ker}(\chi) = \{g \in G \mid \chi(g) = \chi(1_G)\}.$$

Lorsque $K = \mathbb{C}$, c'est un sous-groupe distingué de G, égal au noyau de la représentation dont χ est le caractère associé (cf. lemme 1.7).

On a alors le théorème suivant.

1.10. Théorème. *Soit G un groupe fini. On suppose que $K = \mathbb{C}$. Alors :*

(1) *tout sous-groupe distingué de G est le noyau d'un caractère de G ;*

(2) *tout sous-groupe distingué de G est une intersection finie de noyaux de caractères irréductibles de G.*

Démonstration. D'après le lemme 1.7, le noyau d'un caractère est un sous-groupe distingué de G. Il en est donc de même de toute intersection finie de noyaux de caractères irréductibles.

Inversement, soit H un sous-groupe distingué de G, et soit

$$\rho' : G/H \longrightarrow \mathrm{GL}(V)$$

une représentation fidèle de G/H. On peut prendre par exemple la représentation régulière de G/H par la remarque XXXIV-2.12. Soit $\pi : G \longrightarrow G/H$ la projection canonique, et soit $\rho = \rho' \circ \pi : G \longrightarrow \mathrm{GL}(V)$. Puisque $\mathrm{Ker}(\rho')$ est trivial, on obtient aisément $\mathrm{Ker}(\rho) = H$. Si χ est le caractère de la représentation ρ, le lemme 1.7 donne alors

$$\mathrm{Ker}(\chi) = \mathrm{Ker}(\rho) = H,$$

d'où (1).

Par le théorème de Maschke, $V = V_1 \oplus \cdots \oplus V_r$, où chaque V_i est irréductible. Puisque chaque V_i est stable par G, pour tout $g \in G$, on a

$$\rho(g) = \mathrm{Id}_V \quad \Longleftrightarrow \quad \rho(g)_{|V_i} = \mathrm{Id}_{V_i}, \quad \text{pour tout } i \in [\![1, r]\!].$$

Notons $\rho_i : G \longrightarrow \mathrm{GL}(V_i)$ le morphisme associé à la sous-représentation V_i.

Autrement dit, on a

$$\begin{aligned} \rho_i : G &\longrightarrow \mathrm{GL}(V_i) \\ g &\longmapsto \rho(g)_{|V_i}. \end{aligned}$$

On a donc

$$\mathrm{Ker}(\rho) = \bigcap_{i=1}^{r} \mathrm{Ker}(\rho_i).$$

Si χ_i est le caractère associé à ρ_i, le lemme 1.7 donne alors

$$H = \mathrm{Ker}(\chi) = \bigcap_{i=1}^{r} \mathrm{Ker}(\chi_i).$$

Cela achève la démonstration. □

1.11. Remarque. Ce résultat permet de donner une nouvelle démonstration du théorème 1.8.

2. Orthogonalité des caractères et conséquences

Nous allons maintenant démontrer que les caractères d'un groupe G sont orthogonaux pour une certaine forme bilinéaire. Nous commençons par retraduire le lemme de Schur dans le cas des $K[G]$-modules.

2.1. Lemme (Schur 2.0)

Soit G un groupe, et soient V_1, V_2 deux représentations irréductibles de G sur un corps K.

Alors, tout morphisme de représentations $f : V_1 \longrightarrow V_2$ est soit nul, soit un isomorphisme. En particulier :

(1) *si V_1 et V_2 ne sont pas isomorphes, alors on a $\mathrm{Hom}_G(V_1, V_2) = 0$;*

(2) *si V est irréductible, $\mathrm{End}_G(V)$ est un anneau à division ;*

(3) *si K est algébriquement clos, pour toute représentation irréductible V de G, on a $\mathrm{End}_G(V) = K\mathrm{Id}_V$.*

2.2. Corollaire. *Soit G un groupe fini. On suppose que K est algébriquement clos, et que $\mathrm{car}(K) \nmid |G|$. Soient (V_1, ρ_1) et (V_2, ρ_2) des représentations irréductibles de G, et soit $f \in \mathscr{L}(V_1, V_2)$.*

Alors :

(1) *si V_1 et V_2 ne sont pas isomorphes, on a*

$$\frac{1}{|G|} \sum_{g \in G} \rho_2(g^{-1}) \circ f \circ \rho_1(g) = 0 ;$$

(2) *si* $(V_1,\rho_1)=(V_2,\rho_2)=(V,\rho)$, *il vient* $\mathrm{car}(K)\nmid \dim_K(V)$ *et*

$$\frac{1}{|G|}\sum_{g\in G}\rho(g^{-1})\circ f\circ\rho(g)=\frac{\mathrm{tr}(f)}{\dim_K(V)}\mathrm{Id}_V.$$

Démonstration. Posons $f_G=\dfrac{1}{|G|}\sum\limits_{g\in G}\rho_2(g^{-1})\circ f\circ\rho_1(g)$.

Pour tout $v_1\in V_1$, on a donc

$$f_G(v_1)=\frac{1}{|G|}\sum_{g\in G}g^{-1}\cdot\big(f(g\cdot v_1)\big).$$

Alors, pour tout $g'\in G$ et tout $v_1\in V_1$, on a

$$f_G(g'\cdot v_1)=\frac{1}{|G|}\sum_{g\in G}g^{-1}\cdot\big(f\big(g\cdot(g'\cdot v_1)\big)\big)=\frac{1}{|G|}\sum_{g\in G}g^{-1}\cdot\big(f(gg'\cdot v_1)\big).$$

En faisant le changement de variables $g\longleftrightarrow gg'^{-1}$ et en utilisant la linéarité de l'action de G sur V_2, on obtient

$$f_G(g'\cdot v_1)=\frac{1}{|G|}\sum_{g\in G}g'g^{-1}\cdot\big(f(g\cdot v_1)\big)=g'\cdot f_G(v_1).$$

On a donc $f_G\in\mathrm{Hom}_G(V_1,V_2)$. Par le lemme de Schur, si V_1 et V_2 ne sont pas isomorphes, on a $f_G=0$, d'où (1).

Si $(V_1,\rho_1)=(V_2,\rho_2)=(V,\rho)$, alors $f_G=\lambda_f\mathrm{Id}_V$ pour un certain $\lambda_f\in K$, toujours d'après ce même lemme. On a donc $\mathrm{tr}(f_G)=\dim_K(V)\lambda_f$. Comme on a $\rho(g^{-1})=\rho(g)^{-1}$ pour tout $g\in G$, on a aussi

$$\mathrm{tr}(f_G)=\frac{1}{|G|}\sum_{g\in G}\mathrm{tr}\big(\rho(g^{-1})\circ f\circ\rho(g)\big)=\frac{1}{|G|}\sum_{g\in G}\mathrm{tr}(f)=\mathrm{tr}(f).$$

Ainsi, $\dim_K(V)\lambda_f=\mathrm{tr}(f)$. En appliquant cette égalité à un endomorphisme f de trace égale à 1, on obtient que $\dim_K(V)\cdot 1_K$ est inversible dans K, ce qui revient à dire que $\mathrm{car}(K)\nmid\dim_K(V)$.

En revenant au cas général, on obtient donc que $f_G=\dfrac{\mathrm{tr}(f)}{\dim_K(V)}\mathrm{Id}_V$. Cela achève la démonstration. □

On en déduit le résultat calculatoire suivant.

2.3. Corollaire. *Soit G un groupe fini. On suppose que K est algébriquement clos, et que $\mathrm{car}(K)\nmid|G|$. Soient (V_1,ρ_1), (V_2,ρ_2) et (V,ρ) des représentations irréductibles de G. Soient $\mathbf{e}_1$, $\mathbf{e}_2$ et $\mathbf{e}$ des bases de V_1, V_2*

et V respectivement. Pour tout $g \in G$, on pose

$$\mathrm{Mat}\big(\rho_1(g)\,;\mathbf{e}_1\big) = R_1(g) = \big(r^{(1)}_{i_1 j_1}(g)\big)_{i_1,j_1\in[\![1,n_1]\!]},$$
$$\mathrm{Mat}\big(\rho_2(g)\,;\mathbf{e}_2\big) = R_2(g) = \big(r^{(2)}_{i_2 j_2}(g)\big)_{i_2,j_2\in[\![1,n_2]\!]},$$
$$\mathrm{Mat}\big(\rho(g)\,;\mathbf{e}\big) = R(g) = \big(r_{ij}(g)\big)_{i,j\in[\![1,n]\!]}.$$

On suppose que V_1 et V_2 ne sont pas isomorphes. Alors :

(1) *pour tous i_1, $j_1 \in [\![1,n_1]\!]$ et tous i_2, $j_2 \in [\![1,n_2]\!]$, on a*

$$\frac{1}{|G|}\sum_{g\in G} r^{(2)}_{i_2 j_2}(g^{-1}) r^{(1)}_{i_1 j_1}(g) = 0\,;$$

(2) *pour tous i, j, k, $\ell \in [\![1,n]\!]$, on a*

$$\frac{1}{|G|}\sum_{g\in G} r_{ik}(g^{-1}) r_{\ell j}(g) = \begin{cases} \dfrac{1}{\dim_K(V)}, & \text{si } (i,k)=(j,\ell),\\ 0, & \text{sinon.}\end{cases}$$

Démonstration. Gardons les notations de l'énoncé.

Pour tout $k \in [\![1,n_2]\!]$ et tout $\ell \in [\![1,n_1]\!]$, on a

$$R_2(g^{-1})E_{k\ell}R_1(g) = \big(r^{(2)}_{i_2 k}(g^{-1}) r^{(1)}_{\ell j_1}(g)\big)_{i_2,j_1}.$$

Le corollaire 2.2 (1) se retraduit matriciellement par

$$\frac{1}{|G|}\sum_{g\in G} R_2(g^{-1}) M R_1(g) = 0, \quad \text{pour tout } M \in \mathrm{M}_{n_2\times n_1}(K).$$

En choisissant pour M la matrice élémentaire E_{j_2,i_1}, on obtient (1).

Le point (2) de ce même corollaire se retraduit en revanche par

$$\frac{1}{|G|}\sum_{g\in G} R(g^{-1}) M R(g) = \frac{\mathrm{tr}(M)}{\dim_K(V)}\mathrm{I}_n, \quad \text{pour tout } M \in \mathrm{M}_n(K).$$

En choisissant pour M la matrice élémentaire $E_{k,\ell}$, on obtient

$$\Big(\frac{1}{|G|}\sum_{g\in G} r_{ik}(g^{-1}) r_{\ell j}(g)\Big)_{i,j} = \frac{\mathrm{tr}(E_{k\ell})}{\dim_K(V)}\mathrm{I}_n,$$

soit encore

$$\frac{1}{|G|}\sum_{g\in G} r_{ik}(g^{-1}) r_{\ell j}(g) = \frac{1}{\dim_K(V)}\delta_{i,j}\delta_{k,\ell},$$

d'où (2). □

Notation. On suppose que G est un groupe fini et que $\mathrm{car}(K) \nmid |G|$.

Pour tous $\varphi, \psi \in \mathscr{F}_c(G,K)$, on pose

$$\langle \varphi, \psi \rangle_G = \frac{1}{|G|} \sum_{g \in G} \varphi(g)\psi(g^{-1}).$$

Pour tout $\varphi \in \mathscr{F}_c(G,K)$, on pose également

$$\varphi^*(g) = \varphi(g^{-1}), \quad \text{pour tout } g \in G.$$

Clairement, $\varphi^* \in \mathscr{F}_c(G,K)$.

On a alors le lemme suivant.

2.4. Lemme. *On suppose que G est un groupe fini et que $\mathrm{car}(K) \nmid |G|$. Alors, l'application $\langle\ ,\ \rangle_G : \mathscr{F}_c(G,K) \times \mathscr{F}_c(G,K) \longrightarrow K$ est une forme bilinéaire symétrique.*

De plus, pour tous $C, C' \in \mathrm{Conj}(G)$, on a

$$\langle e_C, e^*_{C'} \rangle_G = \frac{|C|}{|G|} \delta_{C,C'}.$$

Enfin, si $\mathrm{car}(K) = 0$, la forme bilinéaire $\langle\ ,\ \rangle_G$ est non dégénérée.

Démonstration. Le fait que $\langle\ ,\ \rangle_G$ est bilinéaire est clair, et la symétrie est obtenue en effectuant le changement de variables $g \longleftrightarrow g^{-1}$. Considérons la base $(e_C)_{C \in \mathrm{Conj}(G)}$ définie dans le lemme 1.4. Rappelons que pour tout élément $g \in G$ et tout $C \in \mathrm{Conj}(G)$, on a

$$e_C(g) = \begin{cases} 1, & \text{si } g \in C, \\ 0, & \text{sinon.} \end{cases}$$

Pour tous $C, C' \in \mathrm{Conj}(G)$, on a alors

$$\langle e_C, e^*_{C'} \rangle_G = \frac{1}{|G|} \sum_{g \in G} e_C(g) e_{C'}(g) = \frac{|C \cap C'|}{|G|} = \frac{|C|}{|G|} \delta_{C,C'}.$$

Supposons maintenant que $\mathrm{car}(K) = 0$ et soit $\varphi = \sum_C \lambda_C e_C \in \mathscr{F}_c(G,K)$ telle que $\langle \varphi, \psi \rangle_G = 0$, pour tout $\psi \in \mathscr{F}_c(G,K)$. On a donc

$$\langle \varphi, e^*_{C'} \rangle_G = \sum_C \lambda_C \langle e_C, e^*_{C'} \rangle_G = 0, \quad \text{pour tout } C' \in \mathrm{Conj}(G).$$

Le point précédent montre alors que $\dfrac{|C'|}{|G|} \lambda_{C'} = 0$, pour tout $C' \in \mathrm{Conj}(G)$. Puisque $\mathrm{car}(K) = 0$, on en déduit que $\lambda_{C'} = 0$, pour tout $C' \in \mathrm{Conj}(G)$, d'où $\varphi = 0$. Comme $\langle\ ,\ \rangle_G$ est symétrique, cela suffit à démontrer qu'elle est non dégénérée. Cela achève la démonstration. □

On peut maintenant énoncer le résultat clef de cette section.

2.5. Proposition. *Soit G un groupe fini. On suppose que K est algébriquement clos, et que* $\mathrm{car}(K) \nmid |G|$.
Alors :

(1) *pour tout $\chi \in \mathrm{Irr}(G)$, on a $\langle \chi, \chi \rangle_G = 1_K$;*

(2) *pour tous caractères $\chi_1, \chi_2 \in \mathrm{Irr}(G)$ associés à des représentations irréductibles non isomorphes, on a*

$$\langle \chi_1, \chi_2 \rangle_G = 0 ;$$

(3) *deux représentations irréductibles de G de degré fini sont isomorphes si, et seulement si, leurs caractères sont égaux ;*

(4) il y a au plus $|\mathrm{Conj}(G)|$ classes d'isomorphisme de représentations irréductibles de G.

Démonstration. Soient (V_1, ρ_1) et (V_2, ρ_2) deux représentations irréductibles de G de degré fini, de caractères respectifs χ_1 et χ_2. Gardons les notations du corollaire 2.3. On a

$$\begin{aligned} \langle \chi_1, \chi_2 \rangle_G &= \frac{1}{|G|} \sum_{g \in G} \Big(\sum_i r_{ii}^{(1)}(g) \Big) \Big(\sum_j r_{jj}^{(2)}(g^{-1}) \Big) \\ &= \sum_{i,j} \Big(\frac{1}{|G|} \Big(\sum_{g \in G} r_{jj}^{(2)}(g^{-1}) r_{ii}^{(1)}(g) \Big) \Big). \end{aligned}$$

Si les représentations V_1 et V_2 ne sont pas isomorphes, le point (1) du corollaire 2.3 implique alors que $\langle \chi_1, \chi_2 \rangle_G = 0$.

Soit maintenant une représentation irréductible (V, ρ) de caractère χ. Comme précédemment, on a

$$\langle \chi, \chi \rangle_G = \sum_{i,j} \Big(\frac{1}{|G|} \Big(\sum_{g \in G} r_{jj}(g^{-1}) r_{ii}(g) \Big) \Big).$$

En utilisant le point (2) du corollaire 2.3, on obtient

$$\langle \chi, \chi \rangle_G = \sum_{i=1}^{n} \Big(\frac{1}{|G|} \Big(\sum_{g \in G} r_{ii}(g^{-1}) r_{ii}(g) \Big) \Big) = \sum_{i=1}^{n} \frac{1}{\dim_K(V)} = 1_K,$$

d'où (1) et (2).

Supposons maintenant que $\chi_1 = \chi_2$. Par (1), on a

$$\langle \chi_1, \chi_2 \rangle_G = \langle \chi_1, \chi_1 \rangle_G = 1_K.$$

Si V_1 et V_2 étaient non isomorphes, on aurait aussi $\langle \chi_1, \chi_2 \rangle_G = 0$, d'où une contradiction. Ainsi, V_1 et V_2 sont isomorphes, d'où (3).

Il reste à démontrer (4). Montrons que les éléments de $\mathrm{Irr}_K(G)$ forment une famille libre de $\mathscr{F}_c(G,K)$. Soient $\chi_1,\dots,\chi_m$ des éléments de $\mathrm{Irr}_K(G)$ deux à deux distincts. Supposons qu'il existe $\lambda_1,\dots,\lambda_m \in K$ tels que

$$\lambda_1\chi_1 + \cdots + \lambda_m\chi_m = 0.$$

Soit $j \in [\![1,m]\!]$. En utilisant les points (1) et (2), on obtient alors

$$0 = \Big\langle \sum_{i=1}^{m} \lambda_i\chi_i, \chi_j \Big\rangle_G = \sum_{i=1}^{m} \lambda_i \langle \chi_i, \chi_j\rangle_G = \lambda_j 1_K = \lambda_j.$$

Ainsi, les éléments de $\mathrm{Irr}_K(G)$ sont bien K-linéairement indépendants. Comme $\mathscr{F}_c(G,K)$ est de dimension $|\mathrm{Conj}(G)|$ par le lemme 1.4, on en déduit que $\mathrm{Irr}_K(G)$ possède au plus $|\mathrm{Conj}(G)|$ éléments. Le point (3) permet alors de conclure. Cela achève la démonstration. □

Notation. Si V est une représentation de G, et si $m \geqslant 0$ est un entier, on note mV la représentation $V \times \cdots \times V$.

On a alors le corollaire suivant.

2.6. Corollaire. *Soit G un groupe fini. On suppose que K est algébriquement clos, et que $\mathrm{car}(K) \nmid |G|$. Soit $\{V_1,\dots,V_s\}$ un système complet de représentations irréductibles de G (i.e. un ensemble $\{V_1,\dots,V_s\}$ de représentations irréductibles telles que chaque représentation irréductible de G est isomorphe à un unique V_i).*

Soient V et V' deux représentations de G de degré fini, de caractères χ et χ' respectivement. Écrivons

$$V \simeq \prod_{i=1}^{s} m_i V_i \quad \text{et} \quad V' \simeq \prod_{i=1}^{s} m_i' V_i,$$

où m_i, $m_i' \geqslant 0$. Alors, on a

$$\langle \chi, \chi'\rangle_G = \Big(\sum_{i,j} m_i m_j'\Big)\cdot 1_K.$$

Démonstration. Si χ_i est le caractère irréductible associé à V_i, on a

$$\chi = \sum_{i=1}^{s} m_i\chi_i \quad \text{et} \quad \chi' = \sum_{i=1}^{s} m_i'\chi_i$$

d'après l'exemple 1.5(7). Il suffit alors d'appliquer la bilinéarité de l'application $\langle\ ,\ \rangle_G$ et les points (1) et (2) de la proposition précédente. □

On déduit de tout cela un premier résultat important.

2.7. Théorème. *Soit G un groupe fini. On suppose que K est algébriquement clos, et que $\mathrm{car}(K) \nmid |G|$. Soit $\{V_1, \ldots, V_s\}$ un système complet de représentations irréductibles de G (i.e. un ensemble $\{V_1, \ldots, V_s\}$ de représentations irréductibles telles que chaque représentation irréductible de G est isomorphe à un unique V_i).*

Soit V une représentation de G de degré fini, de caractère χ, et écrivons

$$V \simeq \prod_{i=1}^{s} m_i V_i,$$

où $m_i \geqslant 0$, et où $m_i V_i$ est la représentation $V_i \times \cdots \times V_i$, m_i fois.

Si χ_i est le caractère de V_i, on a $\chi = \sum_{i=1}^{s} m_i \chi_i$. De plus,

$$\langle \chi, \chi_i \rangle_G = m_i \cdot 1_K, \quad \textit{pour tout } i \in [\![1, s]\!] \quad \textit{et} \quad \langle \chi, \chi \rangle_G = \Big(\sum_{i=1}^{s} m_i^2 \Big) \cdot 1_K.$$

Si $\mathrm{car}(K) = 0$, on a également les propriétés suivantes :

(1) *la multiplicité m_i de V_i dans la décomposition de V est égale*[1] *à $\langle \chi, \chi_i \rangle_G$. En particulier, m_i est indépendant du choix de la décomposition de V ;*

(2) *deux représentations de G de degré fini sont isomorphes si, et seulement si, leurs caractères sont égaux ;*

(3) *un caractère χ de G est irréductible si, et seulement si, $\langle \chi, \chi \rangle_G = 1_K$.*

Démonstration. La première partie découle du corollaire précédent, appliqué à $V' = V_i$ dans le premier cas et $V' = V$ dans le second.

Supposons maintenant que $\mathrm{car}(K) = 0$. Le point (1) découle du fait que l'application canonique $\mathbb{Z} \longrightarrow K$ est injective, et que l'égalité

$$\langle \chi, \chi_i \rangle_G \;=\; m_i \cdot 1_K$$

détermine m_i de manière unique.

Le point (2) découle alors immédiatement du point (1). Enfin, soit χ le caractère d'une représentation V de degré fini. Si χ est irréductible, par la proposition précédente, on a $\langle \chi, \chi \rangle_G = 1_K$. Inversement, si $\langle \chi, \chi \rangle_G = 1_K$, on a

$$(m_1^2 + \cdots + m_s^2) \cdot 1_K \;=\; 1_K,$$

où m_i est la multiplicité de V_i dans la décomposition de V.

Puisque $\mathrm{car}(K) = 0$, on en déduit que $m_1^2 + \cdots + m_s^2 = 1$. Il existe donc un unique entier m_i égal à 1, et les autres entiers m_j sont nuls. On a alors un isomorphisme de $V \simeq V_i$, et V est irréductible. □

1. On identifie ici $\mathbb{Z}$ avec $\mathbb{Z} \cdot 1_K$, ce qui est licite puisqu'en caractéristique nulle, le morphisme canonique $\mathbb{Z} \longrightarrow K$ est injectif.

Le point (2) de ce théorème n'est plus vrai si $\mathrm{car}(K) > 0$, comme le montre l'exemple suivant.

2.8. Exemple. Soit K un corps algébriquement clos de caractéristique 2 et soit $j \in K$ une racine 3-ième primitive de l'unité. Alors, les représentations

$$\begin{array}{ccc} \rho_1 : \mathbb{Z}/3\mathbb{Z} & \longrightarrow & \mathrm{GL}_3(K) \\ \overline{m} & \longmapsto & j^m \mathrm{I}_3 \end{array} \quad \text{et} \quad \begin{array}{ccc} \rho_2 : \mathbb{Z}/3\mathbb{Z} & \longrightarrow & \mathrm{GL}_3(K) \\ \overline{m} & \longmapsto & \begin{pmatrix} 1 & & \\ & 1 & \\ & & j^m \end{pmatrix} \end{array}$$

ont même caractère, mais ne sont pas isomorphes, puisque le sous-espace des vecteurs de K^3 fixes sous l'action de G est nul dans le premier cas, et de dimension 2 dans le second cas.

2.9. Corollaire. *Soit G un groupe fini. On suppose que K est algébriquement clos, et que $\mathrm{car}(K) \nmid |G|$. Soit V une représentation de G de degré fini, de caractère χ. Alors, on a*

$$\langle \chi, \chi_0 \rangle_G = \dim_K \left(V^G\right) \cdot 1_K.$$

Démonstration. D'après l'exemple XXXIV-4.3 (2), on a

$$V = V^G \oplus V_G,$$

avec $V_G = \left\{ v \in V \,\middle|\, \sum_{g \in G} g \cdot v = 0 \right\}$.

Calculons la multiplicité de la représentation triviale. Tout d'abord,

$$V^G \simeq \dim_K \left(V^G\right) \mathbf{1}.$$

Supposons maintenant que la décomposition de V_G en somme directe de sous-représentations irréductibles contienne au moins un facteur V' isomorphe à la représentation triviale. Alors, G agit trivialement sur V', et ainsi $V' \subset V^G$, ce qui contredit le fait que $V^G \cap V_G = 0$. Ainsi, la multiplicité recherchée est exactement $\dim_K \left(V^G\right)$. On applique alors le théorème précédent pour conclure. □

2.10. Remarque

Le point (3) du théorème 2.7 est faux si K n'est pas algébriquement clos (cf. exercice 7). Néanmoins, puisque les représentations V et $V \otimes_K \overline{K}$ ont même caractère, on déduit le fait suivant.

Soit χ le caractère d'une représentation de G sur K. Alors, $\langle \chi, \chi \rangle_G = 1_K$ si, et seulement si, $V \otimes_K \overline{K}$ est irréductible.

Le résultat suivant montre que le fait de supposer que K est algébriquement clos et de caractéristique 0 n'est pas une hypothèse très contraignante du point de vue du problème de classification des représentations à isomorphisme près.

2.11. Proposition. *Soit K un corps infini, et soit L/K une extension. Alors, deux représentations V et V' d'un groupe fini G sont isomorphes si, et seulement si, les représentations $V \otimes_K L$ et $V' \otimes_K L$ sont isomorphes.*

Démonstration. Notons $\rho : G \longrightarrow \mathrm{GL}(V)$ et $\rho' : G \longrightarrow \mathrm{GL}(V')$ les morphismes de groupes correspondants. Supposons que V et V' soient isomorphes. Il existe donc un isomorphisme de K-espaces vectoriels

$$u : V \xrightarrow{\sim} V'$$

tel que

$$u \circ \rho(g) \circ u^{-1} = \rho'(g), \quad \text{pour tout } g \in G.$$

Par le lemme XXXIV-3.5, $u_L : V \otimes_K L \longrightarrow V' \otimes_K L$ est un isomorphisme de L-espaces vectoriels et

$$u_L \circ \rho(g)_L \circ u_L^{-1} = \rho'(g)_L, \quad \text{pour tout } g \in G,$$

ou encore

$$u_L \circ \rho_L(g) \circ u_L^{-1} = \rho'_L(g), \quad \text{pour tout } g \in G.$$

Ainsi, $V \otimes_K L$ et $V' \otimes_K L$ sont isomorphes.

Supposons maintenant que $V \otimes_K L$ et $V' \otimes_K L$ soient isomorphes, et soit

$$\varphi : V \otimes_K L \xrightarrow{\sim} V' \otimes_K L$$

un isomorphisme de représentations. Soit $\mathbf{e}$ une base de V et soit $\mathbf{e}'$ une base de V'. Pour tout g, on pose

$$M(g) = \mathrm{Mat}\big(\rho(g)\,;\mathbf{e}\big) \quad \text{et} \quad M'(g) = \mathrm{Mat}\big(\rho(g)\,;\mathbf{e}'\big).$$

Si $Q = \mathrm{Mat}(\varphi\,;\mathbf{e}\otimes\mathbf{1},\mathbf{e}'\otimes\mathbf{1}) \in \mathrm{GL}_n(L)$, l'hypothèse se retraduit par

$$QM(g)Q^{-1} = M'(g), \quad \text{pour tout } g \in G.$$

La proposition XXXI-4.10 montre alors qu'il existe $P \in \mathrm{GL}_n(K)$ telle que

$$PM(g)P^{-1} = M'(g), \quad \text{pour tout } g \in G.$$

Ainsi, les représentations V et V' sont isomorphes. Cela achève la démonstration. □

On en déduit une généralisation du point (2) du théorème 2.7.

2.12. Corollaire. *Soit G un groupe fini et soit K un corps de caractéristique nulle. Alors, deux représentations de G de degré fini sont isomorphes si, et seulement si, leurs caractères sont égaux.*

Démonstration. Soient V et V' deux représentations de G, de caractères respectifs χ et χ'. Puisqu'un corps de caractéristique nulle est infini, V et V' sont isomorphes si, et seulement si, $V \otimes_K \overline{K}$ et $V' \otimes_K \overline{K}$ le sont par la proposition précédente. Par le théorème 2.7, ces représentations sont isomorphes si, et seulement si, leurs caractères sont égaux. Mais ces caractères sont respectivement égaux à χ et χ' d'après le dernier point de l'exemple 1.5, d'où le résultat. □

Avertissement ! Dans tout ce qui suit, on suppose que G est un groupe fini et que K est un corps algébriquement clos de caractéristique nulle, sauf mention expresse du contraire. En particulier, on identifie $\mathbb{Z}$ à son image dans K.

Continuons à étudier les représentations de permutation.

2.13. Exemple. Soit $K[E]$ une représentation de permutation de G, où E est fini non vide, et soit χ_E son caractère. On rappelle que

$$\dim_K \left(K[E]^G\right) = |\Omega|,$$

où Ω est l'ensemble des G-orbites de E, et que

$$\chi_E(g) = |\mathrm{Fix}(g)|, \quad \text{pour tout } g \in G.$$

Calculons $\langle \chi_E, \chi_E \rangle_G$.

Remarquons tout d'abord que pour tout $g \in G$, on a

$$\chi_E(g^{-1}) = |\mathrm{Fix}(g^{-1})| = |\mathrm{Fix}(g)| = \chi_E(g),$$

puisque pour tout $x \in E$, on a $g^{-1}\cdot x = x$ si, et seulement si, $g\cdot x = x$. On en déduit alors immédiatement que

$$\langle \chi_E, \chi_E \rangle_G = \langle \chi_E^2, \chi_0 \rangle_G.$$

De plus, le groupe G agit sur $E \times E$ par

$$g\cdot(x, y) = (g\cdot x, g\cdot y), \quad \text{pour tout } g \in G \text{ et tous } x, y \in E.$$

Pour $g \in G$ fixé, et pour tous $x, y \in E$, on a $g\cdot(x, y) = (x, y)$ si, et seulement si, $g\cdot x = x$ et $g\cdot y = y$. Par conséquent, le fixateur de $g \in G$ dans $E \times E$ est l'ensemble $\mathrm{Fix}(g) \times \mathrm{Fix}(g)$, d'où l'égalité

$$\chi_{E\times E} = \chi_E^2.$$

Le corollaire 2.9 montre alors que

$$\langle \chi_E, \chi_E \rangle_G = |\Omega_{E\times E}|,$$

où $\Omega_{E\times E}$ est l'ensemble des G-orbites de $E \times E$.

En explicitant $\langle \chi_E, \chi_E \rangle_G$, on obtient aussi

$$\frac{1}{|G|} \sum_{g\in G} |\mathrm{Fix}(g)|^2 = |\Omega_{E\times E}|.$$

Soit maintenant χ'_E le caractère de $K[E]_G$. On a donc

$$\chi_E = |\Omega|\chi_0 + \chi'_E.$$

De plus, le corollaire 2.9 nous indique que $\langle \chi'_E, \chi_0 \rangle_G = 0$, puisque l'on a $\langle \chi_E, \chi_0 \rangle_G = |\Omega|$.

On a alors

$$|\Omega_{E\times E}| = \langle \chi_E, \chi_E \rangle_G = |\Omega|^2 + \langle \chi'_E, \chi'_E \rangle_G,$$

soit encore $\langle \chi'_E, \chi'_E \rangle_G = |\Omega_{E\times E}| - |\Omega|^2$.

Supposons que G agisse transitivement sur E et que $|E| \geqslant 2$. On a donc

$$\langle \chi'_E, \chi'_E \rangle_G = |\Omega_{E\times E}| - 1.$$

Par conséquent, le théorème 2.7 montre alors que χ'_E est irréductible si, et seulement si, $E \times E$ ne possède que deux G-orbites. Or, la diagonale

$$\Delta = \{(x, x) \in E \times E \mid x \in E\}$$

est une telle G-orbite, puisque G agit transitivement sur E. Ainsi, comme l'ensemble $E \times E$ est réunion disjointe de ses orbites, $(E \times E) \setminus \Delta$ est la réunion disjointe des G-orbites de $E \times E$ distinctes de Δ. Par conséquent, $E \times E$ ne possède que deux G-orbites si, et seulement si, $(E \times E) \setminus \Delta$ est une G-orbite. Cela revient à dire que l'action de G sur E est 2-transitive.

On a donc démontré le résultat suivant. On suppose que G agit transitivement et que $|E| \geqslant 2$. Alors, les propriétés suivantes sont équivalentes :

(1) le caractère χ'_E est irréductible ;

(2) on a $\langle \chi_E^2, \chi_0 \rangle_G = 2$ (ou de manière équivalente, $\langle \chi_E, \chi_E \rangle_G = 2$) ;

(3) l'action de G sur E est 2-transitive.

Si l'on applique ces résultats à l'action de $\mathfrak{S}_n$ sur $E = [\![1, n]\!]$, avec $n \geqslant 3$, on retrouve le fait que $H_n = K[E]_{\mathfrak{S}_n}$ est irréductible, sous l'hypothèse que le corps K est algébriquement clos de caractéristique 0.

On obtient aussi l'égalité

$$\sum_{\sigma\in\mathfrak{S}_n} |\mathrm{Fix}(\sigma)|^2 = 2\cdot n!\,.$$

Nous pouvons maintenant décomposer la décomposition régulière en somme de représentations irréductibles.

2.14. Proposition. *Soit $\{V_1,\dots,V_s\}$ un système complet de représentations irréductibles de G. Alors, on a*

$$K[G] \simeq \dim_K(V_1)V_1 \times \cdots \times \dim_K(V_s)V_s.$$

Démonstration. Soit χ_i le caractère de V_i et soit m_i la multiplicité de V_i dans la décomposition de $K[G]$. D'après le théorème 2.7, on a

$$m_i = \langle \chi_{\mathrm{reg}}, \chi_i\rangle_G = \frac{1}{|G|}\left(|G|\chi_i(1_G^{-1})\right) = \chi_i(1_G) = \dim_K(V_i),$$

d'où le résultat. □

2.15. Corollaire. *Soit $\{V_1,\dots,V_s\}$ un système complet de représentations irréductibles de G, soit m_i le degré de V_i et soit χ_i le caractère de V_i. Alors :*

(1) *on a $m_1^2+\cdots+m_s^2 = |G|$;*

(2) *pour tout $g\in G\setminus\{1_G\}$, $m_1\chi_1(g)+\cdots+m_s\chi_s(g)=0$.*

Démonstration. Puisque $K[G]$ est de dimension $|G|$, le point (1) provient de la proposition précédente et d'une comparaison des dimensions. La proposition précédente montre également que l'on a

$$\chi_{\mathrm{reg}} = m_1\chi_1+\cdots+m_s\chi_s.$$

Il suffit alors d'évaluer en $g\in G\setminus\{1_G\}$ pour obtenir (2). □

Le but de ce qui suit est de déterminer le nombre de classes d'isomorphisme de représentations irréductibles. On commence par un lemme.

2.16. Lemme. *Soit (V,ρ) une représentation irréductible de G, de caractère χ. Alors, pour tout $\varphi\in\mathscr{F}_c(G,K)$, on a*

$$\frac{1}{|G|}\sum_{g\in G}\varphi(g)\rho(g) = \frac{\langle\varphi,\chi^*\rangle_G}{\dim_K(V)}\,\mathrm{Id}_V.$$

Démonstration. Posons $\rho_\varphi = \dfrac{1}{|G|}\displaystyle\sum_{g\in G}\varphi(g)\rho(g)$, et montrons que ρ_φ est un morphisme de représentations.

Pour tout $g' \in G$, et tout $v \in V$, on a

$$\rho_\varphi(g'\cdot v) = \frac{1}{|G|}\sum_{g\in G}\varphi(g)\big(g\cdot(g'\cdot v)\big) = \frac{1}{|G|}\sum_{g\in G}\varphi(g)(gg'\cdot v).$$

En effectuant le changement de variables $g \longleftrightarrow g'gg'^{-1}$, et en utilisant le fait que φ est une fonction centrale, on obtient

$$\rho_\varphi(g'\cdot v) = \frac{1}{|G|}\sum_{g\in G}\varphi(g)\big(g'\cdot(g\cdot v)\big) = g'\cdot\rho_\varphi(v),$$

la dernière égalité utilisant la linéarité de l'action de G. Ainsi, $\rho_\varphi \in \mathrm{End}_G(V)$. Par le lemme de Schur, il existe $\lambda \in K$ tel que $\rho_\varphi = \lambda \mathrm{Id}_V$. En prenant la trace, on obtient

$$\frac{1}{|G|}\sum_{g\in G}\varphi(g)\chi(g) = \langle\varphi,\chi^*\rangle_G = \dim_K(V)\lambda.$$

Le corps K étant de caractéristique nulle, il s'ensuit que $\lambda = \dfrac{\langle\varphi,\chi^*\rangle_G}{\dim_K(V)}$, d'où le résultat voulu. □

On peut maintenant en déduire le résultat suivant.

2.17. Théorème. *Les éléments de* $\mathrm{Irr}(G)$ *forment une base orthonormée de* $\mathscr{F}_c(G,K)$ *pour la forme bilinéaire* $\langle\ ,\ \rangle_G$.

En particulier, le nombre de classes d'isomorphisme de représentations irréductibles et le nombre de caractères irréductibles de G *sont tous deux égaux au nombre de classes de conjugaison de* G.

Démonstration. La proposition 2.5 montre que les éléments de $\mathrm{Irr}(G)$ forment une famille orthonormée. En particulier, cette famille est K-libre. Montrons qu'elle est génératrice.

Soit $F = \mathrm{Vect}_K\big(\chi, \chi \in \mathrm{Irr}(G)\big)$. Puisque $\langle\ ,\ \rangle_G$ est non dégénérée par le lemme 2.4, on a clairement $F = \mathscr{F}_c(G,K)$ si, et seulement si, $F^\perp = 0$, par le théorème XXXII-4.5. Soit $\varphi \in F^\perp$. Remarquons que pour tout $\chi \in \mathrm{Irr}(G)$, on a $\chi^* \in \mathrm{Irr}(G) \subset F$. On a donc en particulier

$$\langle\varphi,\chi^*\rangle_G = 0, \quad \text{pour tout } \chi \in \mathrm{Irr}(G).$$

Le lemme 2.16 montre alors que pour toute représentation irréductible (V,ρ), on a

$$\frac{1}{|G|}\sum_{g\in G}\varphi(g)\rho(g) = 0.$$

Par le théorème de Maschke, on en déduit que cette égalité est vraie pour toute représentation de G de degré fini.

Appliquons cela à la représentation régulière.
Dans ce cas, on a $\rho(g)(e_1) = e_g$, et ainsi

$$\frac{1}{|G|}\sum_{g\in G}\varphi(g)e_g = 0.$$

On en déduit alors que $\varphi(g) = 0$ pour tout $g \in G$, i.e. $\varphi = 0$. On obtient ainsi la première partie du théorème.

On a donc

$$\dim_K\left(\mathscr{F}_c(G,K)\right) = |\mathrm{Irr}(G)|.$$

Or, cette dimension est aussi égale au nombre de classes de conjugaison du groupe G par le lemme 1.4. Enfin, deux représentations irréductibles sont isomorphes si, et seulement si, leurs caractères sont égaux par le théorème 2.7, d'où la seconde partie. □

Profitons-en pour en déduire une caractérisation des groupes abéliens.

2.18. Proposition. *Le groupe G est abélien si, et seulement si, toute représentation irréductible de G est de degré 1.*

Démonstration. Un groupe G est abélien si, et seulement si, toutes ses classes de conjugaison sont des singletons. Lorsque G est fini, cela revient à dire que son nombre s de classes de conjugaison est égal à $|G|$. D'après le corollaire 2.15, on en déduit que G est abélien si, et seulement si

$$s = m_1^2 + \cdots + m_s^2,$$

en tenant compte du fait que le nombre de classes d'isomorphisme de représentations irréductibles est aussi égal à s.

Comme $m_i \geqslant 1$ pour tout $i \in [\![1, s]\!]$, cela revient à dire que $m_i = 1$ pour tout $i \in [\![1, s]\!]$, d'où le résultat. □

On continue en établissant d'autres relations utiles entre les valeurs des caractères.

2.19. Corollaire. *Soient $\chi_1, \ldots, \chi_s$ les caractères irréductibles du groupe G. Soit $g \in G$ et soit C la classe de conjugaison de g. Pour tout $g' \in G$, on a*

$$\frac{1}{|G|}\sum_{i=1}^{s}\chi_i(g')\chi_i(g^{-1}) = \begin{cases} \dfrac{1}{|C|}, & \text{si } g' \in C, \\ 0, & \text{sinon.} \end{cases}$$

Démonstration. Soit $g \in G$, et soit C la classe de conjugaison de g. Puisque e_C est une fonction centrale et que $(\chi_1, \ldots, \chi_s)$ forment une base orthonormée de $\mathscr{F}_c(G,K)$, on a $e_C = \sum_{i=1}^{s}\langle e_C, \chi_i\rangle\chi_i$.

Or, on a $e_C(h) = 1$ si h est conjugué à g et $e_C(h) = 0$ sinon. On a donc

$$\langle e_C, \chi_i \rangle = \frac{1}{|G|} \sum_{h \in C} \chi_i(h^{-1}).$$

Comme χ_i est constante sur la classe de conjugaison de g^{-1}, on a

$$\chi_i(h^{-1}) = \chi_i(g^{-1}), \quad \text{pour tout } h \in C.$$

Ainsi,

$$\langle e_C, \chi_i \rangle = \frac{|C|}{|G|} \chi_i(g^{-1}),$$

soit encore

$$e_C = \frac{|C|}{|G|} \sum_{i=1}^{s} \chi_i(g^{-1}) \chi_i.$$

On en déduit alors le résultat voulu en évaluant cette égalité en g', et en divisant par $|C|$. □

Pour finir cette section, nous allons nous intéresser à un résultat de théorie de groupes, mais dont la seule démonstration connue utilise la théorie des représentations. Le théorème que nous avons en vue est le suivant.

2.20. Théorème (Frobenius). *Soit G un groupe fini, et soit H un sous-groupe de G tel que*

$$H \cap gHg^{-1} = \{1_G\}, \quad \textit{pour tout } g \in G \setminus H.$$

Alors, $N = \Big(G \setminus \bigcup_{g \in G} gHg^{-1}\Big) \cup \{1_G\}$ est un sous-groupe distingué de G, et on a $G = N \rtimes H$.

Démonstration. La démonstration qui suit est une version extrêmement détaillée de la démonstration proposée dans l'annexe A de [44].

Nous allons découper la démonstration en plusieurs points. Nous démontrerons successivement les faits suivants :

(1) Soit $g_1 = 1_G, g_2, \ldots, g_r \in G$ un système complet de représentants de l'ensemble G/H. Alors, on a une réunion disjointe

$$\Big(\bigcup_{g \in G} gHg^{-1}\Big) \setminus \{1_G\} = \bigcup_{i=1}^{r} \big(g_i H g_i^{-1} \setminus \{1_G\}\big).$$

En particulier, $|N| = [G : H]$.

(2) Pour toute fonction centrale $\psi \in \mathscr{F}_c(H, \mathbb{C})$, il existe une unique fonction centrale $\widetilde{\psi} \in \mathscr{F}_c(G, \mathbb{C})$ telle que

$$\widetilde{\psi}_{|H} = \psi \quad \text{et} \quad \widetilde{\psi}(g) = \psi(1_G), \quad \text{pour tout } g \in N.$$

De plus, l'application

$$\begin{aligned} \tau : \mathscr{F}_c(H, \mathbb{C}) &\longrightarrow \mathscr{F}_c(G, \mathbb{C}) \\ \psi &\longmapsto \widetilde{\psi} \end{aligned}$$

est un morphisme de $\mathbb{C}$-algèbres, et l'on a

$$\widetilde{\psi^*} = (\widetilde{\psi})^*, \quad \text{pour tout } \psi \in \mathscr{F}_c(H, \mathbb{C}).$$

(3) Pour tout $\psi \in \mathscr{F}_c(H, \mathbb{C})$, et tout $\varphi \in \mathscr{F}_c(G, \mathbb{C})$, on a l'égalité

$$\langle \widetilde{\psi}, \varphi \rangle_G = \langle \psi, \varphi_{|H} \rangle_H + \psi(1_G)\langle \chi_{0,G}, \varphi \rangle_G - \psi(1_G)\langle \chi_{0,H}, \varphi_{|H} \rangle_H.$$

Ici, $\chi_{0,H}$ et $\chi_{0,G}$ désigne les caractères triviaux de H et G respectivement.

(4) Pour tous ψ_1, $\psi_2 \in \mathscr{F}_c(H, \mathbb{C})$, on a

$$\langle \widetilde{\psi}_1, \widetilde{\psi}_2 \rangle_G = \langle \psi_1, \psi_2 \rangle_H.$$

(5) Si χ est un caractère de H, alors $\widetilde{\chi}$ est un caractère de G.

Montrons comment les points (1)−(5) permettent d'en déduire le théorème. Soit (V, ρ) une représentation fidèle de H, par exemple, la représentation régulière (cf. remarque XXXIV-2.12), et soit χ son caractère. D'après le point (5), il existe une représentation $(\widetilde{V}, \widetilde{\rho})$ de caractère $\widetilde{\chi}$, où $\widetilde{\chi}$ est défini comme dans le point (2).

Nous allons maintenant démontrer que $\mathrm{Ker}(\widetilde{\rho}) = N$. Remarquons tout d'abord que l'on a $\widetilde{\chi}(1_G) = \chi(1_G)$, puisque $1_G \in N$. Ainsi, on a

$$\mathrm{Ker}(\widetilde{\rho}) = \mathrm{Ker}(\widetilde{\chi}) = \{g' \in G \mid \widetilde{\chi}(g') = \chi(1_G)\}$$

d'après le lemme 1.7 (3). Notons maintenant que G est la réunion disjointe de N et de $\big(\bigcup_{g \in G} gHg^{-1}\big) \setminus \{1_G\}$, par définition.

Soit $g' \in G$. Si $g' \in \big(\bigcup_{g \in G} gHg^{-1}\big) \setminus \{1_G\}$, il existe $g \in G$ et $h \in H$ tels que $g' = ghg^{-1}$. Notons que $h \neq 1_G$ (car sinon, on aurait $g' = gg^{-1} = 1_G$). Mais alors, on a

$$\widetilde{\chi}(g') = \widetilde{\chi}(h) = \chi(h),$$

la première égalité découlant du fait qu'un caractère est une fonction centrale, et la seconde découlant de la définition de $\widetilde{\chi}$. Comme ρ est une représentation fidèle, on a $\mathrm{Ker}(\chi) = \mathrm{Ker}(\rho) = \{1_G\}$. Puisque $h \neq 1_G$, on obtient ainsi $\widetilde{\chi}(g') \neq \chi(1_G)$, et $g' \notin \mathrm{Ker}(\widetilde{\rho})$.

Si maintenant $g' \in N$, par définition de $\widetilde{\chi}$, on a $\widetilde{\chi}(g') = \chi(1_G)$, et par conséquent $g' \in \mathrm{Ker}(\widetilde{\rho})$. Finalement, $\mathrm{Ker}(\widetilde{\rho}) = N$, ce qui démontre que N est un sous-groupe distingué de G.

Il reste à démontrer que $G = N \rtimes H$. On sait déjà que N est un sous-groupe distingué de G. De plus, par définition même, $N \cap H = \{1_G\}$. Enfin, par le point (1), on a

$$|N| \times |H| = [G : H] \times |H| = |G|.$$

On applique alors le corollaire X-3.11 pour conclure.

Passons à la démonstration des cinq points précédents.

(1) Soit $g \in G$. Par choix des éléments $g_1, \ldots, g_r$, il existe $i \in [\![1, r]\!]$, et $h \in H$ tels que $g = g_i h$. On a alors

$$gHg^{-1} = (g_i h)H(g_i h)^{-1} = g_i(hHh^{-1})g_i^{-1} = g_i H g_i^{-1}.$$

Cela démontre déjà l'égalité

$$\Big(\bigcup_{g\in G} gHg^{-1}\Big) \setminus \{1_G\} = \bigcup_{i=1}^{r} (g_i H g_i^{-1} \setminus \{1_G\}).$$

Montrons que la réunion est disjointe. Soient $i, j \in [\![1, r]\!]$ tels que

$$\big((g_i H g_i^{-1}) \setminus \{1_G\}\big) \cap \big((g_j H g_j^{-1}) \setminus \{1_G\}\big) \neq \emptyset.$$

Soient $h, h' \in H$ tels que $g_i h g_i^{-1} = g_j h' g_j^{-1}$. Comme ces deux éléments sont distincts de 1_G par hypothèse, il en est de même de h'. Or, on a

$$(g_j^{-1} g_i) h (g_j^{-1} g_i)^{-1} = h'.$$

Comme $h' \neq 1_G$, l'hypothèse sur H entraîne que $g_j g_i^{-1} \in H$, soit

$$\overline{g}_i = \overline{g}_j \in G/H.$$

On en déduit que $i = j$. Ainsi, la réunion précédente est disjointe.

On obtient ainsi une réunion disjointe

$$\bigcup_{g\in G} gHg^{-1} = \bigcup_{i=1}^{r} (g_i H g_i^{-1} \setminus \{1_G\}) \cup \{1_G\}.$$

Comme la conjugaison par un élément de G est un automorphisme, on a alors

$$\begin{aligned}
\Big|\bigcup_{g\in G} gHg^{-1}\Big| &= 1 + \sum_{i=1}^{r} |g_i H g_i^{-1} \setminus \{1_G\})| \\
&= 1 + \sum_{i=1}^{r} (|g_i H g_i^{-1}| - 1) \\
&= 1 + \sum_{i=1}^{r} (|H| - 1) \\
&= 1 + r(|H| - 1) \\
&= 1 + [G : H](|H| - 1) \\
&= 1 + |G| - [G : H].
\end{aligned}$$

On en déduit alors

$$|N| = |G| - (1 + |G| - [G : H]) + 1 = [G : H],$$

d'où (1).

(2) Soit $\psi \in \mathscr{F}_c(H, \mathbb{C})$. Remarquons que, par définition, G est la réunion disjointe de N et de $(\bigcup_{g\in G} gHg^{-1}) \setminus \{1_G\}$.

Supposons qu'une fonction $\widetilde{\psi}$ vérifiant les propriétés de l'énoncé existe. En utilisant le fait que $\widetilde{\psi}$ est centrale, on a

$$\widetilde{\psi}(g') = \psi(1_G), \quad \text{pour tout } g' \in N,$$

ainsi que

$$\widetilde{\psi}(ghg^{-1}) = \psi(h), \quad \text{pour tout } g \in G \text{ et tout } h \in H.$$

La remarque précédente montre alors qu'il y a au plus une telle fonction $\widetilde{\psi}$.

Montrons maintenant que l'application $\widetilde{\psi}$ donnée par ces égalités est bien définie. Si $ghg^{-1} = \tilde{g}\tilde{h}\tilde{g}^{-1} \neq 1_G$, avec $g, \tilde{g} \in G$ et $h, \tilde{h} \in H$, il faut montrer que $\psi(h) = \psi(\tilde{h})$. Or, on a

$$(g^{-1}\tilde{g})\tilde{h}(g^{-1}\tilde{g})^{-1} = h \in H.$$

Si $g^{-1}\tilde{g} \in G \setminus H$, l'hypothèse sur H implique que $h = 1_G$, d'où la contradiction $\tilde{g}h\tilde{g}^{-1} = 1_G$. Ainsi, $g^{-1}\tilde{g} \in H$ et par conséquent, il existe $h' \in H$ tel que $\tilde{g} = gh'$. On obtient ainsi

$$ghg^{-1} = g(h'\tilde{h}h'^{-1})g^{-1},$$

d'où $h = h'\tilde{h}h'^{-1}$. Comme ψ est une fonction centrale, on obtient alors l'égalité $\psi(h) = \psi(\tilde{h})$, d'où le résultat. Notons que cette dernière égalité est encore vraie lorsque $ghg^{-1} = 1_G$, puisqu'alors $h = \tilde{h} = 1_G$.

Par construction, on a ainsi

$$\widetilde{\psi}(g') = \psi(1_G), \quad \text{pour tout } g' \in N \quad \text{et} \quad \widetilde{\psi}(h) = \psi(h), \quad \text{pour tout } h \in H.$$

Il reste à vérifier que $\widetilde{\psi}$ est une fonction centrale. Notons que $\bigcup_{g \in G} gHg^{-1}$ et $\{1_G\}$ sont stables par conjugaison. Il en est donc de même de N. Pour tout $g' \in N$, et tout $g \in G$, on a alors

$$\widetilde{\psi}(gg'g^{-1}) = \psi(1_G) = \widetilde{\psi}(g').$$

Si maintenant $g, g' \in G$ et $h \in H$, on a

$$\widetilde{\psi}\big(g'(ghg^{-1})g'^{-1}\big) = \widetilde{\psi}\big((g'g)h(g'g)^{-1}\big) = \psi(h) = \widetilde{\psi}(ghg^{-1}).$$

Montrons enfin que l'application

$$\begin{array}{rccc} \tau : & \mathscr{F}_c(H, \mathbb{C}) & \longrightarrow & \mathscr{F}_c(G, \mathbb{C}) \\ & \psi & \longmapsto & \widetilde{\psi} \end{array}$$

est un morphisme de $\mathbb{C}$-algèbres. Soient $\psi_1, \psi_2, \psi \in \mathscr{F}_c(H, \mathbb{C})$. Alors, on a

$$(\widetilde{\psi_1} + \widetilde{\psi_2})_{|H} = (\widetilde{\psi_1})_{|H} + (\widetilde{\psi_2})_{|H} = \psi_1 + \psi_2.$$

De plus, pour tout $g' \in N$, on a

$$(\widetilde{\psi_1} + \widetilde{\psi_2})(g') = \psi_1(1_G) + \psi_2(1_G) = (\psi_1 + \psi_2)(1_G).$$

Par définition de $\widetilde{\psi_1 + \psi_2}$, on obtient

$$\widetilde{\psi_1 + \psi_2} = \widetilde{\psi_1} + \widetilde{\psi_2}.$$

On montre de même que

$$\widetilde{\psi_1\psi_2} = \widetilde{\psi_1}\widetilde{\psi_2} \quad \text{et} \quad \widetilde{\lambda\psi} = \lambda\widetilde{\psi}, \quad \text{pour tout } \lambda \in \mathbb{C},$$

ainsi que $\widetilde{\psi^*} = (\widetilde{\psi})^*$. Cela démontre (2).

(3) Commençons par remarquer que la formule à vérifier est vraie dans le cas où $\psi = \chi_{0,H}$, puisque l'on a alors $\widetilde{\psi} = \chi_{0,G}$. En constatant que

$$\psi = \big(\psi - \psi(1_G)\chi_{0,H}\big) + \psi(1_G)\chi_{0,H},$$

et utilisant la $\mathbb{C}$-linéarité de l'application τ, on voit alors que l'on peut se réduire au cas où $\psi(1_G) = 0$. Dans ce cas, la formule à démontrer se récrit

$$\langle\widetilde{\psi}, \varphi\rangle_G = \langle\psi, \varphi_{|H}\rangle_H,$$

pour tout $\psi \in \mathscr{F}_c(H, \mathbb{C})$, et tout $\varphi \in \mathscr{F}_c(G, \mathbb{C})$.

Soit $\psi \in \mathscr{F}_c(H, \mathbb{C})$ tel que $\psi(1_G) = 0$. On a donc $\widetilde{\psi}(g) = 0$, pour tout $g \in N$. En utilisant le point (1), on obtient

$$\langle\widetilde{\psi}, \varphi\rangle_G = \frac{1}{|G|} \sum_{g \in G\setminus N} \widetilde{\psi}(g)\varphi(g^{-1}),$$

pour tout $\varphi \in \mathscr{F}_c(G, \mathbb{C})$. Comme tout élément de $G \setminus N$ s'écrit de manière unique sous la forme $g_i h g_i^{-1}$, avec $i \in [\![1, r]\!]$ et $h \in H \setminus \{1_G\}$, en tenant compte du fait que $\widetilde{\psi}(1_G)$ est nul, on a ainsi

$$\langle\widetilde{\psi}, \varphi\rangle_G = \frac{1}{|G|} \sum_{i=1}^{r} \sum_{h \in H} \widetilde{\psi}(g_i h g_i^{-1})\varphi(g_i^{-1}h^{-1}g_i).$$

Puisque $\widetilde{\psi}$ et φ sont des fonctions centrales, on a

$$\begin{aligned}
\langle\widetilde{\psi}, \varphi\rangle_G &= \frac{1}{|G|} \sum_{i=1}^{r} \sum_{h \in H} \widetilde{\psi}(h)\varphi(h^{-1}) \\
&= \frac{[G : H]}{|G|} \sum_{h \in H} \widetilde{\psi}(h)\varphi(h^{-1}) \\
&= \frac{1}{|H|} \sum_{h \in H} \psi(h)\varphi(h^{-1}) \\
&= \langle\psi, \varphi_{|H}\rangle_H,
\end{aligned}$$

d'où (3).

(4) Une application du point (3) montre que pour tout $\psi \in \mathscr{F}_c(H, \mathbb{C})$, on a

$$\langle \widetilde{\psi}, \chi_{0,G} \rangle_G = \langle \psi, \chi_{0,H} \rangle_H.$$

Mais alors, pour tous ψ_1, $\psi_2 \in \mathscr{F}_c(H, \mathbb{C})$, on a

$$\begin{aligned} \langle \widetilde{\psi}_1, \widetilde{\psi}_2 \rangle_G &= \langle \widetilde{\psi}_1 (\widetilde{\psi}_2)^*, \chi_{0,G} \rangle_G \\ &= \langle \widetilde{\psi_1 \psi_2^*}, \chi_{0,G} \rangle_G \\ &= \langle \psi_1 \psi_2^*, \chi_{0,H} \rangle_H \\ &= \langle \psi_1, \psi_2 \rangle_H. \end{aligned}$$

Cela démontre (4).

(5) Exceptionnellement, notons $\theta_1, \ldots, \theta_s$ les caractères irréductibles de G. Soit χ un caractère de H. Puisque $\widetilde{\chi}$ est une fonction centrale, et que la famille $(\theta_1, \ldots, \theta_s)$ est une base orthonormée de $\mathscr{F}_c(G, \mathbb{C})$ pour la forme bilinéaire $\langle\ ,\rangle_G$ (cf. théorème 2.17), on a

$$\widetilde{\chi} = \sum_{i=1}^{s} \langle \widetilde{\chi}, \theta_i \rangle_G \theta_i.$$

Remarquons maintenant que le point (3) et le corollaire 2.6 entraînent que $c_i = \langle \widetilde{\chi}, \theta_i \rangle_G \in \mathbb{Z}$, pour tout $i \in [\![1, s]\!]$. De plus, on a

$$\langle \widetilde{\chi}, \widetilde{\chi} \rangle_G = \sum_{i=1}^{s} c_i^2.$$

Supposons pour commencer que χ est irréductible. On va alors montrer que $\widetilde{\chi}$ est un caractère irréductible de G.

Par le point (4), on obtient

$$\langle \widetilde{\chi}, \widetilde{\chi} \rangle_G = \sum_{i=1}^{s} c_i^2 = \langle \chi, \chi \rangle_H = 1.$$

Comme $c_1, \ldots, c_s \in \mathbb{Z}$, il existe un unique entier $k \in [\![1, s]\!]$ tel que $c_k^2 = 1$, et $c_i = 0$, pour tout $i \neq k$. On a donc $\widetilde{\chi} = \pm \theta_k$. Si $\widetilde{\chi} = -\theta_k$, alors

$$\widetilde{\chi}(1_G) = \chi(1_G) = -\theta_k(1_G) < 0,$$

ce qui est impossible puisque $\chi(1_G) \geqslant 1$. Finalement, $\widetilde{\chi} = \theta_k$, ce qui est le résultat annoncé.

Supposons maintenant que χ soit un caractère de H arbitraire. On sait alors que $\chi = m_1 \chi_1 + \cdots + m_t \chi_t$, où $\chi_1, \ldots, \chi_t$ sont les caractères irréductibles de H, et $m_1, \ldots, m_t$ sont des entiers positifs. Par (2), on a

$$\widetilde{\chi} = m_1 \widetilde{\chi}_1 + \cdots + m_t \widetilde{\chi}_t.$$

D'après ce qui précède, $\widetilde{\chi}_1, \ldots, \widetilde{\chi}_t$ sont des caractères irréductibles de G.

Si $\widetilde{V}_i$ est une représentation irréductible dont $\widetilde{\chi}_i$ est le caractère, on en déduit que $\widetilde{\chi}$ est le caractère de la représentation $m_1\widetilde{V}_1 \times \cdots \times m_t\widetilde{V}_t$, d'où le point (5). Cela achève la démonstration. □

3. Table des caractères d'un groupe fini

Nous allons maintenant nous concentrer sur le cas où $K = \mathbb{C}$. Le lemme qui suit nous permettra de réinterpréter plus simplement les divers résultats obtenus sur les valeurs des caractères irréductibles.

3.1. Lemme. *Soit $K = \mathbb{C}$ et soit G un groupe fini. Pour tout caractère χ de G, et tout $g \in G$, on a*

$$\chi(g^{-1}) = \overline{\chi(g)}.$$

Démonstration. Soit (V, ρ) une représentation de degré fini de G. Soit l'élément $g \in G$, et soit $d = o(g)$. D'après le lemme 1.7 (1), $\rho(g)$ est diagonalisable et, ses valeurs propres sont des racines d-ièmes de l'unité. Il existe donc une base $\mathbf{e}$ de V telle que

$$\mathrm{Mat}\big(\rho(g);\mathbf{e}\big) = \begin{pmatrix} z_1 & & \\ & \ddots & \\ & & z_n \end{pmatrix},$$

où $z_1, \ldots, z_n$ sont des racines d-ièmes de l'unité. On a donc $z_i^{-1} = \overline{z}_i$ pour tout $i \in [\![1, n]\!]$. Mais alors, puisque $\rho(g^{-1}) = \rho(g)^{-1}$, on a

$$\mathrm{Mat}\big(\rho(g^{-1});\mathbf{e}\big) = \begin{pmatrix} z_1 & & \\ & \ddots & \\ & & z_n \end{pmatrix}^{-1} = \begin{pmatrix} z_1^{-1} & & \\ & \ddots & \\ & & z_n^{-1} \end{pmatrix} = \begin{pmatrix} \overline{z}_1 & & \\ & \ddots & \\ & & \overline{z}_n \end{pmatrix}.$$

Mais alors, on a $\chi(g) = z_1 + \cdots + z_n$, ainsi que

$$\chi(g^{-1}) \;=\; \overline{z}_1 + \cdots + \overline{z}_n \;=\; \overline{\chi(g)}.$$

Cela achève la démonstration. □

On a vu que les caractères déterminaient les représentations de G (de degré fini) à isomorphisme près. De plus, le théorème de Maschke implique que les caractères des représentations de G de degré fini sont exactement les combinaisons linéaires à coefficients positifs des caractères irréductibles. Il est donc important de déterminer les valeurs des caractères irréductibles et d'en garder trace.

Notation. Soit G un groupe fini. Si $\chi \in \mathrm{Irr}(G)$ et C est une classe de conjugaison de G, on note $\chi(C)$ l'élément $\chi(g) \in K$, où g est un élément arbitraire de C.

Dans la suite, on note $C_1, \ldots, C_s$ les diverses classes de conjugaison de G et $\chi_1, \ldots, \chi_s$ les divers caractères irréductibles de G.

On définit alors un tableau de taille $s \times s$, donc la i-ème ligne est

$$\big(\chi_i(C_1), \ldots, \chi_i(C_s)\big).$$

Ce tableau est appelé *la table de caractères de G.*

Compte tenu du lemme 3.1, les divers résultats sur les valeurs des caractères irréductibles se retraduisent alors comme suit.

Pour tous $i, j \in [\![1, s]\!]$, on a

(1) $\displaystyle\sum_{k=1}^{s} |C_k| \chi_i(C_k) \overline{\chi_j(C_k)} = \delta_{ij} |G|$;

(2) $\displaystyle\sum_{k=1}^{s} \chi_k(C_i) \overline{\chi_k(C_j)} = \delta_{ij} \frac{|G|}{|C_i|}$.

Si maintenant $X = (x_1, \ldots, x_s)$ et $Y = (y_1, \ldots, y_s)$ sont deux vecteurs lignes à coefficients dans $\mathbb{C}$, on pose

$$(X \mid Y)_{G,\ell} = \sum_{k=1}^{s} |C_k| x_k \overline{y}_k.$$

Si $X = \begin{pmatrix} x_1 \\ \vdots \\ x_s \end{pmatrix}$ et $Y = \begin{pmatrix} y_1 \\ \vdots \\ y_s \end{pmatrix}$ sont cette fois des vecteurs colonnes, on pose

$$(X \mid Y)_{G,c} = \sum_{k=1}^{s} x_k \overline{y}_k.$$

Si $\mathscr{L}_1, \ldots, \mathscr{L}_s$ sont les lignes de la table de caractères de G, et si $\mathscr{C}_1, \ldots, \mathscr{C}_s$ en sont les colonnes, les relations précédentes se récrivent alors ainsi.

Pour tous $i, j \in [\![1, s]\!]$, on a

(1) $(\mathscr{L}_i \mid \mathscr{L}_j)_{G,\ell} = \delta_{ij} |G|$;

(2) $(\mathscr{C}_i \mid \mathscr{C}_j)_{G,c} = \delta_{ij} \dfrac{|G|}{|C_i|}$.

Les lignes et les colonnes de la table de caractères de G vérifient donc certaines relations d'orthogonalité, bien pratiques pour compléter la table lorsqu'il manque quelques valeurs.

Donnons maintenant quelques exemples. Pour faciliter la lecture de la table de caractères, on écrira un représentant de chaque classe de conjugaison au sommet de chaque colonne.

3.2. Exemple. Soit $G = \mathbb{Z}/3\mathbb{Z}$. Le groupe G étant abélien, toutes les classes de conjugaison sont réduites à un élément, et toutes les représentations irréductibles sont de degré 1, d'après la proposition 2.18. De plus, vu l'exemple 1.5 (1), il suffit de trouver tous les morphismes de groupes $G \longrightarrow \mathbb{C}^\times$. Comme G est d'exposant 3, l'image d'un tel morphisme est contenue dans $\mathbb{U}_3$ et entièrement déterminée par l'image de $\overline{1}$. On obtient donc la table de caractères suivante.

$\mathbb{Z}/3\mathbb{Z}$	$\overline{0}$	$\overline{1}$	$\overline{2}$
χ_0	1	1	1
χ_1	1	j	j^2
χ_1'	1	j^2	j

3.3. Exemple. Soit $G = (\mathbb{Z}/2\mathbb{Z})^2$. Comme précédemment, toute représentation irréductible est de degré 1. Comme G est d'exposant 2, l'image d'un tel morphisme est contenue dans $\{\pm 1\}$. On obtient alors facilement la table suivante.

$(\mathbb{Z}/2\mathbb{Z})^2$	$(\overline{0},\overline{0})$	$(\overline{1},\overline{0})$	$(\overline{0},\overline{1})$	$(\overline{1},\overline{1})$
χ_0	1	1	1	1
χ_1	1	-1	1	-1
χ_1'	1	1	-1	-1
χ_1''	1	-1	-1	1

3.4. Exemple. Considérons $G = \mathfrak{S}_3$. Il y a trois classes de conjugaison : l'identité, l'ensemble des transpositions, et l'ensemble des 3-cycles (cf. corollaire VIII-2.25).

On a donc exactement trois classes d'isomorphismes de représentations irréductibles. Or, on en connaît déjà trois : la représentation triviale, celle donnée par la signature ε, et la représentation H_3. Cette dernière a été décrite complètement dans l'exemple XXXIV-2.19 (5).

Les caractères de ces représentations seront notés χ_0, ε et χ_2 respectivement. Les calculs faits dans l'exemple XXXIV-2.19 (5) permettent alors de calculer les valeurs de χ_2. On obtient la table de caractères suivante :

$\mathfrak{S}_3$	Id	(1 2)	(1 2 3)
χ_0	1	1	1
ε	1	-1	1
χ_2	2	0	-1

Notons que l'on peut retrouver le fait que $\mathfrak{S}_3$ possède trois classes de conjugaison, ainsi que cette table de caractères, quasiment sans calculs.

En effet, on sait que $m_1^2 + \cdots + m_s^2 = 6$. Cette égalité montre déjà que l'on a $1 \leqslant m_i \leqslant 2$ pour tout entier $i \in [\![1, s]\!]$, puisque $3^2 > 6$. D'autre part, comme $\mathfrak{S}_3$ est non abélien, il existe alors au moins un des entiers m_i qui est égal à 2, par la proposition 2.18. De plus, il ne peut pas y avoir deux m_i égaux à 2, car $2^2 + 2^2 > 6$. Bref, il existe un unique m_i égal à 2, les autres étant tous égaux à 1.

On en déduit alors que $s = 3$, ainsi que $m_1 = m_2 = 1$ et $m_3 = 2$ (quitte à renuméroter). On a donc deux représentations irréductibles de degré 1 et une représentation irréductible de degré 2 (à isomorphisme près, bien sûr).

Une fois que l'on a choisi un représentant de chaque classe, on remplit la table de caractères avec les valeurs de χ_0 et ε.

$\mathfrak{S}_3$	Id	(1 2)	(1 2 3)
χ_0	1	1	1
ε	1	-1	1
χ_2	?	?	?

On peut alors calculer les valeurs manquantes de χ_2, sans connaître la représentation associée, en utilisant les relations d'orthogonalité sur les colonnes. Sachant que χ_2 est un caractère associé à une représentation de degré 2, on sait déjà que $\chi_2(\mathrm{Id}) = 2$.

$\mathfrak{S}_3$	Id	(1 2)	(1 2 3)
χ_0	1	1	1
ε	1	-1	1
χ_2	2	?	?

On complète alors la table par orthogonalité des colonnes (ou grâce au point (2) du corollaire 2.15).

3.5. Exemple. Soit $G = D_4$. L'égalité $m_1^2 + \cdots + m_s^2 = 8$ montre que chaque m_i est inférieur ou égal à 2. Le groupe G n'étant pas abélien, un des m_i est égal à 2. De plus, il y a aussi au moins un m_i égal à 1, puisque la représentation triviale est irréductible de degré 1. Il ne peut alors y avoir deux représentations irréductibles de degré 2, car $1^2 + 2^2 + 2^2 > 8$. Autrement dit, il existe un unique entier m_i égal à 2, tous les autres étant égaux à 1. En résumé, D_4 possède quatre représentations irréductibles de degré 1, et une de degré 2. En particulier, il y a cinq classes de conjugaison.

On peut d'ailleurs vérifier ce fait directement. Un système complet de représentants des classes de conjugaison est donné par

$$\mathrm{Id}_{\mathbb{R}^2}, \quad -\mathrm{Id}_{\mathbb{R}^2}, \quad r, \quad s, \quad rs,$$

où r est la rotation de centre O et de mesure d'angle $\frac{\pi}{2}$, et s est la symétrie orthogonale d'axe (Ox).

Remarquons que $G/\{\pm \mathrm{Id}_{\mathbb{R}^2}\}$ est isomorphe à $(\mathbb{Z}/2\mathbb{Z})^2$. Plus précisément, on a un morphisme surjectif

$$\begin{array}{rccc} \pi : & G & \longrightarrow & (\mathbb{Z}/2\mathbb{Z})^2 \\ & r^k s^\ell & \longmapsto & (\overline{k}, \overline{\ell}) \end{array}$$

de noyau $\{\pm \mathrm{Id}_{\mathbb{R}^2}\}$. Par inflation, on en déduit quatre représentations irréductibles de degré 1, qui sont exactement les quatre caractères recherchés. Cela est absolument normal d'après la remarque 1.2 (2), puisque l'on peut vérifier que $[G, G] = \{\pm \mathrm{Id}_{\mathbb{R}^2}\}$.

Les éléments $\pm \mathrm{Id}_{\mathbb{R}^2}$ sont envoyés sur $(\overline{0}, \overline{0})$ par π, tandis que r, s et rs sont envoyés sur $(\overline{1}, \overline{0})$, $(\overline{0}, \overline{1})$ et $(\overline{1}, \overline{1})$ respectivement. Grâce à la table de caractères de $(\mathbb{Z}/2\mathbb{Z})^2$, on peut donc commencer à remplir la table de caractères comme suit.

D_4	$\mathrm{Id}_{\mathbb{R}^2}$	$-\mathrm{Id}_{\mathbb{R}^2}$	r	s	rs
χ_0	1	1	1	1	1
χ_1	1	1	−1	1	−1
χ_1'	1	1	1	−1	−1
χ_1''	1	1	−1	−1	1
χ_2	2	?	?	?	?

On la complète ensuite grâce aux relations d'orthogonalité, par exemple, ce qui donne finalement la table suivante.

D_4	$\mathrm{Id}_{\mathbb{R}^2}$	$-\mathrm{Id}_{\mathbb{R}^2}$	r	s	rs
χ_0	1	1	1	1	1
χ_1	1	1	−1	1	−1
χ_1'	1	1	1	−1	−1
χ_1''	1	1	−1	−1	1
χ_2	2	−2	0	0	0

Notons qu'il est aisé de trouver la représentation irréductible de degré 2. C'est simplement l'inclusion $D_4 \subset \mathrm{GL}_2(\mathbb{R}^2)$. Elle est bien irréductible, puisque r ne stabilise aucune droite, et par conséquent il ne peut y avoir de droite stable par tout élément de D_4.

En écrivant les matrices représentatives de r et s dans la base canonique, on obtient d'ailleurs la version matricielle de cette représentation. C'est l'unique morphisme $\rho : D_4 \longrightarrow \mathrm{GL}_2(\mathbb{R})$ tel que

$$\rho(r) = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{et} \quad \rho(s) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

3.6. Exemple. Soit $G = \mathfrak{A}_4$. On vérifie que

$$[G,G] = \langle (1\ 2)(3\ 4),\ (1\ 3)(2\ 4)\rangle \simeq (\mathbb{Z}/2\mathbb{Z})^2,$$

et que l'on a un isomorphisme

$$G/[G,G] \simeq \mathbb{Z}/3\mathbb{Z},$$

envoyant la classe de $(1\ 2\ 3)$ sur $\overline{1}$.

On a donc exactement trois représentations irréductibles de degré 1, toute autre représentation irréductible étant donc de degré au moins 2. Or, la seule décomposition de douze en somme de carrés contenant exactement trois termes égaux à un est

$$12 = 1^2 + 1^2 + 1^2 + 3^2.$$

Il y a donc une dernière représentation irréductible, qui est de degré 3.

On peut aussi retrouver ce fait en constatant directement qu'il y a exactement quatre classes de conjugaisons dans $\mathfrak{A}_4$, représentées par

$$\mathrm{Id}, \quad (1\ 2)(3\ 4), \quad (1\ 2\ 3) \quad \text{et} \quad (1\ 3\ 2).$$

Par inflation des caractères de $\mathbb{Z}/3\mathbb{Z}$, puis en utilisant l'orthogonalité des colonnes, on obtient la table de caractères suivante.

$\mathfrak{A}_4$	Id	$(1\ 2)(3\ 4)$	$(1\ 2\ 3)$	$(1\ 3\ 2)$
χ_0	1	1	1	1
χ_1	1	1	j	j^2
χ_1'	1	1	j^2	j
χ_3	3	-1	0	0

On vérifie sans mal que la restriction de la représentation[2] H_4 de $\mathfrak{S}_4$ au groupe $\mathfrak{A}_4$ a pour caractère χ_3. C'est donc la représentation de degré 3 recherchée. Les calculs sont laissés au lecteur à titre d'exercice.

3.7. Exemple. On considère $G = \mathfrak{S}_5$. Le groupe G possède sept classes de conjugaison $C_1, \ldots, C_7$, représentées respectivement par

$$\sigma_1 = \mathrm{Id}, \quad \sigma_2 = (1\ 2)(3\ 4), \quad \sigma_3 = (1\ 2\ 3), \quad \sigma_4 = (1\ 2\ 3\ 4\ 5)$$
$$\sigma_5 = (1\ 2), \quad \sigma_6 = (1\ 2\ 3\ 4), \quad \sigma_7 = (1\ 2)(3\ 4\ 5).$$

On sait déjà que les seuls morphismes $\mathfrak{S}_5 \longrightarrow \mathbb{C}^\times$ sont le morphisme trivial et la signature. D'autre part, on a une représentation irréductible H_5 de degré 4, dont le caractère χ_4 a été calculé dans l'exemple 1.5 (11).

2. Représentation définie dans l'exemple 1.5 (11).

On a donc le début de la table de caractères.

$\mathfrak{S}_5$	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7
χ_0	1	1	1	1	1	1	1
ε	1	1	1	1	−1	−1	−1
χ_4	4	0	1	−1	2	0	−1

Par ailleurs, d'après l'exemple XXXIV-2.22 (3), la représentation $H_5(\varepsilon)$ est aussi irréductible, et son caractère est $\varepsilon\chi_4$. Or, on a

$$(\varepsilon\chi_4)(\sigma_5) = -\chi_4(\sigma_5) \neq \chi_4(\sigma_5).$$

Ainsi, $H_5(\varepsilon)$ et H_5 ne sont pas isomorphes. On obtient donc une autre ligne de la table de caractères.

$\mathfrak{S}_5$	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7
χ_0	1	1	1	1	1	1	1
ε	1	1	1	1	−1	−1	−1
χ_4	4	0	1	−1	2	0	−1
$\varepsilon\chi_4$	4	0	1	−1	−2	0	1

Nous allons maintenant calculer le caractère $\chi_{4,s}$ de $\mathrm{Sym}^2(H_5)$. Remarquons que $\sigma_1^2, \sigma_2^2, \sigma_5^2 \in C_1$, $\sigma_3^2 \in C_3$, $\sigma_4^2 \in C_4$, $\sigma_6^2 \in C_2$ et $\sigma_7^2 \in C_3$. En utilisant l'exemple 1.5(8), on obtient les valeurs de $\chi_{4,s}$, que nous isolons pour l'instant de la table de caractères, tant que nous ne sommes pas certains de son irréductibilité.

$\mathfrak{S}_5$	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7
χ_0	1	1	1	1	1	1	1
ε	1	1	1	1	−1	−1	−1
χ_4	4	0	1	−1	2	0	−1
$\varepsilon\chi_4$	4	0	1	−1	−2	0	1
$\chi_{4,s}$	10	2	1	0	4	0	1

On vérifie aisément que

$$|C_1|=1,\ |C_2|=15,\ |C_3|=20,\ |C_4|=24,\ |C_5|=10,\ |C_6|=30,\ |C_7|=20,$$

puis que $\langle \chi_{4,s}, \chi_{4,s}\rangle_G = 3$. On a aussi $\langle \chi_{4,s}, \chi_0\rangle_G = 1$ et $\langle \chi_{4,s}, \chi_4\rangle_G = 1$. D'après le théorème 2.7, on en déduit

$$\chi_{4,s} = \chi_0 + \chi_4 + \chi_5,$$

où χ_5 est le caractère d'une représentation de degré 5 de G convenable (à savoir, la somme des sous-représentations irréductibles de $\mathrm{Sym}^2(H_5)$ non isomorphes à $\mathbf{1}$ ou à H_5, comptées avec multiplicités).

De petits calculs laissés au lecteur montrent alors que

$$\langle \chi_5, \chi_5 \rangle_G = 1.$$

Ainsi, $\chi_5 = \chi_{4,s} - \chi_0 - \chi_4$ est irréductible, d'où une nouvelle ligne de la table. Mais on vérifie en comparant les valeurs en σ_5 que $\varepsilon\chi_5 \neq \chi_5$. Ainsi, le caractère $\varepsilon\chi_5$ est aussi irréductible. On a ainsi la table partielle suivante.

$\mathfrak{S}_5$	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7
χ_0	1	1	1	1	1	1	1
ε	1	1	1	1	−1	−1	−1
χ_4	4	0	1	−1	2	0	−1
$\varepsilon\chi_4$	4	0	1	−1	−2	0	1
χ_5	5	1	−1	0	1	−1	1
$\varepsilon\chi_5$	5	1	−1	0	−1	1	−1

Il reste finalement un dernier caractère à trouver. Calculons d'abord son degré d. La relation

$$120 = 1^2 + 1^2 + 4^2 + 4^2 + 5^2 + 5^2 + d^2$$

montre que $d = 6$. On utilise alors les relations d'orthogonalité sur les colonnes pour en déduire la table complète.

$\mathfrak{S}_5$	σ_1	σ_2	σ_3	σ_4	σ_5	σ_6	σ_7
χ_0	1	1	1	1	1	1	1
ε	1	1	1	1	−1	−1	−1
χ_4	4	0	1	−1	2	0	−1
$\varepsilon\chi_4$	4	0	1	−1	−2	0	1
χ_5	5	1	−1	0	1	−1	1
$\varepsilon\chi_5$	5	1	−1	0	−1	1	−1
χ_6	6	−2	0	1	0	0	0

On peut aussi calculer le caractère de $\mathrm{Alt}^2(H_5)$, qui est de degré 6, et vérifier qu'il est irréductible. On retrouve ainsi χ_6.

4. Décomposition en somme de représentations irréductibles

Pour continuer ce chapitre, nous allons donner une méthode explicite pour décomposer une représentation en somme directe de représentations irréductibles. On commence par un lemme.

4.1. Lemme. *Soit K un corps algébriquement clos de caractéristique nulle, soit G un groupe fini et soit $\{V_1, \dots, V_s\}$ un système complet de représentations irréductibles de G. Pour tout $i \in [\![1, s]\!]$, on note χ_i le caractère de V_i. Enfin, soit (V, ρ) une représentation de degré fini. Alors :*

(1) *il existe une décomposition en somme directe de sous-représentations*

$$V = V^{(1)} \oplus \cdots \oplus V^{(s)},$$

où $V^{(i)} \simeq m_i V_i$, pour tout $i \in [\![1, s]\!]$, pour un certain entier[3] *$m_i \geqslant 0$;*

(2) *pour tout $i \in [\![1, s]\!]$, l'application*

$$p_i = \frac{\dim_K(V_i)}{|G|} \sum_{g \in G} \chi_i(g^{-1}) \rho(g)$$

est le projecteur sur $V^{(i)}$ parallèlement à $\bigoplus_{j \neq i} V^{(j)}$.

(3) *la décomposition $V = V^{(1)} \oplus \cdots \oplus V^{(s)}$ est unique, à ordre près des facteurs, et indépendante du choix de $\{V_1, \dots, V_s\}$.*

Démonstration. D'après le théorème de Maschke, V se décompose en somme directe interne d'un nombre fini de sous-représentations irréductibles. Pour tout $i \in [\![1, s]\!]$, on note $V^{(i)}$ la somme directe des sous-représentations qui sont isomorphes à V_i. On obtient alors une décomposition de la forme voulue, d'où (1).

Soit $i \in [\![1, s]\!]$ et soit W une représentation irréductible, de caractère χ. Rappelons le morphisme correspondant à la sous-représentation W est donnée par

$$\begin{aligned} G &\longrightarrow \mathrm{GL}(W) \\ g &\longmapsto \rho(g)_{|W}. \end{aligned}$$

Le lemme 2.16, appliqué à $\varphi = \chi_i^*$, montre alors que l'on a

$$(p_i)_{|W} = \frac{\dim_K(V_i) \langle \chi_i^*, \chi^* \rangle_G}{\dim_K(W)} \mathrm{Id}_W.$$

En faisant le changement de variables $g \longleftrightarrow g^{-1}$ dans le produit scalaire, cela se récrit

$$(p_i)_{|W} = \frac{\dim_K(V_i) \langle \chi_i, \chi \rangle_G}{\dim_K(W)} \mathrm{Id}_W.$$

Par définition de $\{V_1, \dots, V_s\}$, il existe $j \in [\![1, s]\!]$ tel que $W \simeq V_j$. On a donc $\dim_K(W) = \dim_K(V_j)$ et $\chi = \chi_j$. La proposition 2.5 montre alors que $(p_i)_{|W} = 0$ si $j \neq i$ et $(p_i)_{|W} = \mathrm{Id}_W$ si $j = i$.

3. Si $m_i = 0$, on a $V^{(i)} = 0$.

Au vu de la définition de $V^{(1)}, \ldots, V^{(s)}$, on obtient

$$(p_i)_{|_{V^{(j)}}} = \delta_{ij}\mathrm{Id}_{V^{(i)}}, \quad \text{pour tous } i, j \in [\![1, s]\!].$$

Ces égalités caractérisent exactement le projecteur sur $V^{(i)}$ parallèlement au sous-espace $\bigoplus_{j\neq i} V^{(j)}$, d'où (2).

Notons maintenant que la définition de p_i ne dépend que du degré et du caractère de V_i, qui ne dépendent que de la classe d'isomorphisme de V_i. Ainsi, p_i ne dépend pas du choix du système de représentants irréductibles, une fois un ordre choisi, mais seulement du groupe G. Mais alors, le point (2) montre que $V^{(i)} = \mathrm{Im}(p_i)$. En particulier, la décomposition du point (1) ne dépend que du groupe G. Cela achève la démonstration. □

4.2. Définition. Soit K un corps algébriquement clos de caractéristique nulle, soit G un groupe fini et soit V une représentation de G de degré fini. Les sous-représentations $V^{(1)}, \ldots, V^{(s)}$ définies dans le lemme précédent s'appelle les *composantes isotypiques* de V.

Le lemme précédent nous permet de décomposer une représentation en somme de composantes isotypiques. Il reste à comprendre comment décomposer chaque composante isotypique en somme de sous-représentations toutes isomorphes à une même représentation irréductible.

Soit K un corps algébriquement clos de caractéristique nulle, soit G un groupe fini. Choisissons un système complet de représentations irréductibles de G, noté $\{(V_1, \rho_1), \ldots, (V_s, \rho_s)\}$.

On se fixe un entier $i \in [\![1, s]\!]$, et une base $\mathbf{e}^{(i)} = (e_1^{(i)}, \ldots, e_{n_i}^{(i)})$ de V_i. Pour tout $g \in G$, on pose

$$\mathrm{Mat}\big(\rho_i(g)\,;\mathbf{e}^{(i)}\big) = \big(r_{k\ell}^{(i)}(g)\big)_{k,\ell} \in \mathrm{M}_{n_i}(K).$$

On se donne V une représentation de degré fini, et l'on considère

$$V = V^{(1)} \oplus \cdots \oplus V^{(s)}$$

sa décomposition en composantes isotypiques.

On a alors le résultat suivant.

4.3. Théorème. *Pour tous k, $\ell \in [\![1, n_i]\!]$, on pose*

$$p_{k\ell}^{(i)} = \frac{n_i}{|G|} \sum_{g\in G} r_{\ell k}^{(i)}(g^{-1})\rho(g) \in \mathscr{L}(V).$$

Alors :

(1) *pour tout* $k \in [\![1, n_i]\!]$, *l'endomorphisme* $p_{kk}^{(i)}$ *est un projecteur dont la restriction à* $V^{(j)}$ *est nulle pour tout* $j \neq i$. *De plus, son image* $V_k^{(i)}$ *est contenue dans* $V^{(i)}$ *et l'on a*

$$V^{(i)} = \bigoplus_{k=1}^{n_i} V_k^{(i)},$$

et enfin

$$p_i = \sum_{k=1}^{n_i} p_{kk}^{(i)},$$

où p_i *est le projecteur sur* $V^{(i)}$ *parallèlement à* $\bigoplus_{j \neq i} V^{(j)}$;

(2) *l'endomorphisme* $p_{k\ell}^{(i)}$ *est identiquement nul sur* $V^{(j)}$ *pour tout* $j \neq i$, *ainsi que sur* $V_q^{(i)}$ *pour tout* $q \neq \ell$. *Enfin,* $\mathrm{Im}(p_{k\ell}^{(i)}) \subset V_k^{(i)}$ *et* $p_{k\ell}^{(i)}$ *induit un isomorphisme de* $V_\ell^{(i)}$ *sur* $V_k^{(i)}$;

(3) *soit* $x_1 \in V_1^{(i)} \setminus \{0\}$, *et soit* $x_{1,k} = p_{k1}^{(i)}(x_1) \in V_k^{(i)}$. *La famille*

$$(x_{1,1}, \ldots, x_{1,n_i})$$

est K*-libre, et engendre un sous-espace* $V^{(i)}(x_1)$ *stable par* G. *Pour tout* $g \in G$, *on a*

$$\rho(g)(x_{1,\ell}) = \sum_{k=1}^{n_i} r_{k\ell}(g) x_{1,k}.$$

En particulier, la sous-représentation $V^{(i)}(x_1)$ *est isomorphe à* V_i ;

(4) *si* $(x_1, \ldots, x_{m_i})$ *est une base de* $V_1^{(i)}$, *alors on a une décomposition en somme directe interne de sous-représentations*

$$V^{(i)} = V^{(i)}(x_1) \oplus \cdots \oplus V^{(i)}(x_{m_i}).$$

Démonstration. Quelque soit W une sous-représentation irréductible de V, isomorphe à V_j, intervenant dans une décomposition de V en somme directe de représentations irréductibles. Comme W est isomorphe à V_j, il existe une base $\mathbf{f} = (f_1, \ldots, f_{n_j})$ telle que

$$\rho(g)(f_\ell) = \sum_{k=1}^{n_j} r_{k\ell}^{(j)}(g) f_k, \quad \text{pour tout } \ell \in [\![1, n_j]\!].$$

On a alors

$$\begin{aligned} p_{k\ell}^{(i)}(f_t) &= \frac{n_i}{|G|} \sum_{g \in G} r_{\ell k}^{(i)}(g^{-1}) \Big(\sum_{q=1}^{n_j} r_{qt}^{(j)}(g) f_q \Big) \\ &= \frac{n_i}{|G|} \sum_{q=1}^{n_j} \Big(\sum_{g \in G} r_{\ell k}^{(i)}(g^{-1}) r_{qt}^{(j)}(g) \Big) f_q. \end{aligned}$$

Supposons tout d'abord que $j \neq i$. Le corollaire 2.3 donne alors

$$p_{k\ell}^{(i)}(f_t) = 0, \quad \text{pour tout } t \in [\![1, n_j]\!].$$

Autrement dit :

(i) si $W \simeq V_j$ avec $j \neq i$, la restriction de $p_{k\ell}^{(i)}$ à W est nulle.

Supposons maintenant que $j = i$. Ce même corollaire montre alors que

$$p_{k\ell}^{(i)}(f_t) = \begin{cases} f_k, & \text{si } \ell = t \\ 0, & \text{sinon.} \end{cases}$$

Pour tout $t \in [\![1, n_i]\!]$, on a donc

$$\sum_{k=1}^{n_i} p_{kk}^{(i)}(f_t) = \sum_{k=1}^{n_i} \delta_{k,t} f_k = f_t.$$

Par conséquent :

(ii) si $W \simeq V_i$, la restriction de $\sum_{k=1}^{n_i} p_{kk}^{(i)}$ à W est Id_W.

Enfin, pour tous k, ℓ, q, m, $t \in [\![1, n_j]\!]$, on a alors

$$p_{k\ell}^{(i)} \circ p_{qm}^{(i)}(f_t) = p_{k\ell}^{(i)}(\delta_{m,t} f_q) = \delta_{m,t}\delta_{\ell,q} f_k.$$

On en déduit :

(iii) $p_{k\ell}^{(i)} \circ p_{qm}^{(i)} = \begin{cases} p_{km}^{(i)} & \text{si } q = \ell \\ 0 & \text{sinon.} \end{cases}$

Avant de passer à la démonstration proprement dite, nous allons démontrer une dernière égalité, à savoir :

(iv) $\rho(g) \circ p_{qt}^{(i)} = \sum_{k=1}^{n_i} r_{kq}^{(i)}(g) p_{kt}^{(i)}$.

On a

$$\begin{aligned} \sum_{k=1}^{n_i} r_{kq}^{(i)}(g) p_{kt}^{(i)} &= \frac{n_i}{|G|} \sum_{k=1}^{n_i} r_{kq}^{(i)}(g) \Big(\sum_{h \in G} r_{tk}^{(i)}(h^{-1}) \rho(h) \Big) \\ &= \frac{n_i}{|G|} \sum_{h \in G} \Big(\sum_{k=1}^{n_i} r_{tk}^{(i)}(h^{-1}) r_{kq}^{(i)}(g) \Big) \rho(h). \end{aligned}$$

Comme $\mathrm{Mat}\big(\rho(h^{-1})_{|V_i}; \mathbf{e}\big)\mathrm{Mat}\big(\rho(g)_{|V_i}; \mathbf{e}\big) = \mathrm{Mat}\big(\rho(h^{-1}g)_{|V_i}; \mathbf{e}\big)$, on en déduit

$$\begin{aligned} \sum_{k=1}^{n_i} r_{kq}^{(i)}(g) p_{kt}^{(i)} &= \frac{n_i}{|G|} \sum_{h \in G} r_{tq}^{(i)}(h^{-1}g) \rho(h) \\ &= \frac{n_i}{|G|} \sum_{h \in G} r_{tq}^{(i)}(h^{-1}) \rho(gh) \\ &= \frac{n_i}{|G|} \sum_{h \in G} r_{tq}^{(i)}(h^{-1}) \big(\rho(g) \circ \rho(h) \big) \\ &= \rho(g) \Big(\frac{n_i}{|G|} \sum_{h \in G} r_{tq}^{(i)}(h^{-1}) \rho(h) \Big) \\ &= \rho(g) \circ p_{qt}^{(i)}, \end{aligned}$$

d'où le résultat.

Montrons maintenant les points (1) à (4). Le point (iii) montre en particulier que $p_{kk}^{(i)}$ est un projecteur. De plus, le point (i) implique que la restriction de $p_{kk}^{(i)}$ à $V^{(j)}$ est nulle si $j \neq i$. Par conséquent, son image $V_k^{(i)}$ est contenue dans $V^{(i)}$.

Les points (i) et (ii) montrent également que la restriction de $\sum_{k=1}^{n_i} p_{kk}^{(i)}$ à $V^{(j)}$ est nulle si $j \neq i$, et égale à $\mathrm{Id}_{V^{(i)}}$ si $j = i$. Cela établit l'égalité

$$p_i = \sum_{k=1}^{n_i} p_{kk}^{(i)}.$$

Pour démontrer (1), il reste donc à vérifier l'égalité

$$V^{(i)} = \bigoplus_{k=1}^{n_i} V_k^{(i)}.$$

Or, puisque $V^{(i)} = \mathrm{Im}(p_i)$ et $V_k^{(i)} = \mathrm{Im}\big(p_{kk}^{(i)}\big)$, l'égalité précédente montre que l'on a

$$V^{(i)} = \sum_{k=1}^{n_i} V_k^{(i)}.$$

Il reste à montrer que la somme est directe. Supposons que l'on ait une égalité de la forme

$$v_1 + \cdots + v_{n_i} = 0,$$

avec $v_k \in V_k^{(i)}$. Pour tout $k \in [\![1, n_i]\!]$, il existe donc un vecteur $v'_k \in V$ tel que $v_k = p_{kk}^{(i)}(v'_k)$, et l'on a

$$\sum_{k=1}^{n_i} p_{kk}^{(i)}(v'_k) = 0.$$

En appliquant $p_{\ell\ell}^{(i)}$ à cette égalité et en utilisant (iii), on obtient que

$$v_\ell = p_{\ell\ell}^{(i)}(v'_\ell) = 0, \qquad \text{pour tout } \ell \in [\![1, n_i]\!],$$

d'où le résultat voulu.

Le point (i) montre immédiatement que la restriction de $p_{k\ell}^{(i)}$ à $V^{(j)}$ est nulle si $j \neq i$. Le point (iii) montre aussi que sa restriction à $\mathrm{Im}\big(p_{qq}^{(i)}\big) = V_q^{(i)}$ est nulle pour tout $q \neq \ell$.

Démontrons maintenant que $p_{k\ell}^{(i)}(V^{(i)}) \subset V_k^{(i)}$ pour tous $k, \ell \in [\![1, n_i]\!]$. Cela, joint au fait précédent, montrera que $\mathrm{Im}\big(p_{k\ell}^{(i)}\big) \subset V_k^{(i)}$.

Le point (1) montre que la composante d'un vecteur $v \in V^{(i)}$ sur $V_q^{(i)}$ est $p_{qq}^{(i)}(v)$. Or, on a $p_{qq}^{(i)} \circ p_{k\ell}^{(i)} = \delta_{q,k} p_{q\ell}^{(i)}$. En particulier, la composante du vecteur $p_{k\ell}^{(i)}(v)$ sur $V_q^{(i)}$ est nulle pour tout $q \neq k$. La décomposition en somme directe établie en (1) montre alors que

$$\mathrm{Im}\big(p_{k\ell}^{(i)}\big) \subset V_k^{(i)}.$$

Montrons maintenant que $p_{k\ell}^{(i)}$ induit un isomorphisme de $V_\ell^{(i)}$ sur $V_k^{(i)}$. Soit $v \in V_\ell^{(i)}$ tel que $p_{k\ell}^{(i)}(v) = 0$. Alors, on a

$$0 = p_{\ell k}^{(i)}\big(p_{k\ell}^{(i)}(v)\big) = p_{\ell\ell}^{(i)}(v) = v,$$

la dernier égalité venant du fait que $p_{\ell\ell}^{(i)}$ est un projecteur sur $V_\ell^{(i)}$.

Ainsi, l'application

$$\begin{aligned} V_\ell^{(i)} &\longrightarrow V_k^{(i)} \\ v &\longmapsto p_{k\ell}^{(i)}(v) \end{aligned}$$

est injective. Si maintenant $v \in V_k^{(i)}$, on a

$$v = p_{kk}^{(i)}(v) = p_{k\ell}^{(i)}\big(p_{\ell k}^{(i)}(v)\big).$$

Or, $p_{\ell k}^{(i)}(v) \in \mathrm{Im}(p_{\ell k}^{(i)}) \subset V_\ell^{(i)}$, d'où la surjectivité de cette application. Cela démontre (2).

Montrons (3). Vérifions que $(x_{1,1}, \ldots, x_{1,n_i})$ est libre. Supposons qu'il existe des scalaires $\lambda_1, \ldots, \lambda_{n_i} \in K$ tels que

$$\sum_{k=1}^{n_i} \lambda_k x_{1,k} = 0.$$

Comme $\lambda_k x_{1,k} \in V_k^{(i)}$, le fait que $V_1^{(i)}, \ldots, V_{n_i}^{(i)}$ soient en somme directe montre alors que

$$\lambda_k x_{1,k} = 0, \quad \text{pour tout } k \in [\![1, n_i]\!].$$

Comme x_1 est non nul, et que la restriction de $p_{k1}^{(i)}$ à $V_1^{(i)}$ est injective par (2), on en déduit que $x_{1,k} \neq 0$ pour tout $k \in [\![1, n_i]\!]$. Cela implique alors que $\lambda_k = 0$ pour tout $k \in [\![1, n_i]\!]$. Ainsi, la famille $(x_{1,1}, \ldots, x_{1,n_i})$ est libre. En particulier, $V^{(i)}(x_1)$ est un sous-espace de dimension n_i.

Par (iv), pour tout $\ell \in [\![1, n_i]\!]$, on a

$$\rho(g)(x_{1,\ell}) = \big(\rho(g) \circ p_{\ell 1}^{(i)}\big)(x_1) = \sum_{k=1}^{n_i} r_{k\ell}^{(i)}(g) p_{k1}^{(i)}(x_1) = \sum_{q=1}^{n_i} r_{k\ell}^{(i)}(g) x_{1,k}.$$

Cela montre que $V^{(i)}(x_1)$ est G-stable, et que $V^{(i)}(x_1) \simeq V_i$.

Il reste à démontrer (4). Soit $(x_1, \dots, x_{m_i})$ une K-base de $V_1^{(i)}$. Pour tout entier $j \in [\![1, m_i]\!]$, et pour tout entier $k \in [\![1, n_i]\!]$, on pose $x_{j,k} = p_{k1}^{(i)}(x_j)$. Ainsi, la famille $(x_{j,1}, \dots, x_{j,n_i})$ est une base de $V^{(i)}(x_j)$.

Puisque $p_{k1}^{(i)}$ induit un isomorphisme de $V_1^{(i)}$ sur $V_k^{(i)}$ par (2), la famille

$$(x_{1,k}, \dots, x_{m_i,k})$$

est une base de $V_k^{(i)}$. Par (1), on en déduit que la famille

$$(x_{1,1}, \dots, x_{m_i,1}, \dots, x_{1,n_i}, \dots, x_{m_i,n_i})$$

est une base de $V^{(i)}$. En permutant les termes, on constate que la famille

$$(x_{1,1}, \dots, x_{1,n_i}, \dots, x_{m_i,1}, \dots, x_{m_i,n_i})$$

est une base de $V^{(i)}$. Mais cette famille est exactement celle obtenue en recollant les bases de $V^{(i)}(x_1), \dots, V^{(i)}(x_{m_i})$ précédentes. On en déduit la décomposition en somme directe voulue. Cela achève la démonstration [4].□

En résumé, pour décomposer V en somme de sous-représentations irréductibles, on procède donc comme suit.

On choisit un système complet de représentations irréductibles de G, que l'on notera $\{(V_1, \rho_1), \dots, (V_s, \rho_s)\}$.

On se fixe un entier $i \in [\![1, s]\!]$, et une base $\mathbf{e}^{(i)} = (e_1^{(i)}, \dots, e_{n_i}^{(i)})$ de V_i. Pour tout $g \in G$, on pose

$$\mathrm{Mat}\big(\rho_i(g)\,;\mathbf{e}^{(i)}\big) = \big(r_{k\ell}^{(i)}(g)\big)_{k,\ell} \in \mathrm{M}_{n_i}(K).$$

Pour tout $k \in [\![1, n_i]\!]$, on pose

$$p_{k1}^{(i)} = \frac{n_i}{|G|} \sum_{g \in G} r_{1k}^{(i)}(g^{-1})\rho(g) \in \mathscr{L}(V).$$

Pour chaque $i \in [\![1, s]\!]$, on choisit une base $(x_1, \dots, x_{m_i})$ de $V_1^{(i)} = \mathrm{Im}(p_{11}^{(i)})$. On considère alors

$$V^{(i)}(x_j) = \mathrm{Vect}_K\big(p_{11}^{(i)}(x_j), \dots, p_{n_i1}^{(i)}(x_j)\big).$$

Alors, chaque sous-représentation $V^{(i)}(x_j)$ est isomorphe à V_i, et l'on a

$$V^{(i)} = \bigoplus_{j=1}^{m_i} V^{(i)}(x_j), \quad \text{ainsi que} \quad V = \bigoplus_{i=1}^{s} \bigoplus_{j=1}^{m_i} V^{(i)}(x_j).$$

4. La démonstration présentée ici est une version très détaillée de celle proposée dans [45].

Notons au passage que la dimension m_i de $V_1^{(i)}$ est la multiplicité de V_i dans la décomposition de V.

Enfin, la composante isotypique $V^{(i)}$ est l'image de

$$p_i = \frac{\dim_K(V_i)}{|G|} \sum_{g \in G} \chi_i(g^{-1})\rho(g),$$

où χ_i est le caractère de V_i.

4.4. Exemple. Soit $\mathbb{C}[\mathfrak{S}_3]$ la représentation régulière de $\mathfrak{S}_3$. On note ρ le morphisme correspondant. On pose $c = (1\ 2\ 3)$ et $t = (2\ 3)$. On note V_1 la représentation triviale, V_2 la représentation donnée par la signature.

On a donc $n_1 = n_2 = 1$, ainsi que

$$r_{11}^{(1)}(c) = 1, \quad r_{11}^{(1)}(t) = 1,$$

ainsi que

$$r_{11}^{(2)}(c) = 1, \quad r_{11}^{(2)}(t) = -1.$$

L'exemple XXXIV-2.19 (5) fournit une représentation matricielle V_3 isomorphe à H_3, donnée par

$$\rho_3(c) = \begin{pmatrix} 0 & 1 \\ -1 & -1 \end{pmatrix} \quad \text{et} \quad \rho_3(t) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

On a alors

$$\rho_3(c^2) = \begin{pmatrix} -1 & -1 \\ 1 & 0 \end{pmatrix}, \quad \rho_3(ct) = \begin{pmatrix} 1 & 0 \\ -1 & -1 \end{pmatrix} \quad \text{et} \quad \rho_3(c^2t) = \begin{pmatrix} -1 & -1 \\ 0 & 1 \end{pmatrix}.$$

Ainsi, on a

$$p_{11}^{(1)} = \frac{1}{6} \sum_{g \in \mathfrak{S}_3} \rho(g).$$

La multiplicité de V_1 dans $\mathbb{C}[\mathfrak{S}_3]$ étant égale à 1, on sait que $V_1^{(1)} = \mathrm{Im}(p_{11}^{(1)})$ est de dimension 1. Comme

$$p_{11}^{(1)}(e_{\mathrm{Id}}) = \frac{1}{6} \sum_{g \in \mathfrak{S}_3} e_g \neq 0,$$

on obtient

$$V_1^{(1)} = \mathbb{C}\Big(\sum_{g \in \mathfrak{S}_3} e_g\Big).$$

De même, on a

$$p_{11}^{(2)} = \frac{1}{6}\left(\mathrm{Id} + \rho(c) + \rho(c^2) - \rho(t) - \rho(ct) - \rho(c^2t)\right).$$

Cet endomorphisme est aussi de rang 1, et l'on obtient de manière similaire

$$V_1^{(2)} = \mathbb{C}(e_{\mathrm{Id}} + e_c + e_{c^2} - e_t - e_{ct} - e_{c^2t}).$$

On a également $$p_{11}^{(3)} = \frac{1}{3}\big(\mathrm{Id} - \rho(c) + \rho(ct) - \rho(c^2t)\big),$$

ainsi que $$p_{21}^{(3)} = \frac{1}{3}\big(-\rho(c) + \rho(c^2) + \rho(t) - \rho(c^2t)\big).$$

Ici, $V_1^{(3)}$ est de dimension 2. Or, on a

$$p_{11}^{(3)}(e_{\mathrm{Id}}) = \frac{1}{3}(e_{\mathrm{Id}} - e_c + e_{ct} - e_{c^2t}) \quad \text{et} \quad p_{11}^{(3)}(e_c) = \frac{1}{3}(e_c - e_{c^2} + e_t - e_{ct}).$$

Les deux éléments $x_1 = e_{\mathrm{Id}} - e_c + e_{ct} - e_{c^2t}$ et $x_2 = e_c - e_{c^2} + e_t - e_{ct}$ de $V_1^{(3)}$ sont clairement linéairement indépendants, et (x_1, x_2) est donc une base de $V_1^{(3)}$.

De simples calculs montrent alors que

$$p_{21}^{(3)}(x_1) = -e_c + e_{c^2} + e_t - e_{c^2t} \quad \text{et} \quad p_{21}^{(3)}(x_2) = e_{\mathrm{Id}} - e_{c^2} - e_{ct} + e_{c^2t}.$$

Bref, si l'on pose

$$\begin{aligned} f_1 &= e_{\mathrm{Id}} + e_c + e_{c^2} + e_t + e_{ct} + e_{c^2t} \\ f_2 &= e_{\mathrm{Id}} + e_c + e_{c^2} - e_t - e_{ct} - e_{c^2t} \\ f_3 &= e_{\mathrm{Id}} - e_c + e_{ct} - e_{c^2t} \\ f_4 &= -e_c + e_{c^2} + e_t - e_{c^2t} \\ f_5 &= e_c - e_{c^2} + e_t - e_{ct} \\ f_6 &= e_{\mathrm{Id}} - e_{c^2} - e_{ct} + e_{c^2t}, \end{aligned}$$

on a la décomposition

$$V = \mathrm{Vect}_{\mathbb{C}}(f_1) \oplus \mathrm{Vect}_{\mathbb{C}}(f_2) \oplus \mathrm{Vect}_{\mathbb{C}}(f_3, f_4) \oplus \mathrm{Vect}_{\mathbb{C}}(f_5, f_6),$$

avec $\mathrm{Vect}_{\mathbb{C}}(f_1) \simeq \mathbf{1}$, $\mathrm{Vect}_{\mathbb{C}}(f_2) \simeq \mathbb{C}(\varepsilon)$, et

$$\mathrm{Vect}_{\mathbb{C}}(f_3, f_4) \simeq \mathrm{Vect}_{\mathbb{C}}(f_5, f_6) \simeq H_3.$$

5. Propriétés d'intégralité et applications

Nous allons maintenant démontrer des propriétés d'intégralité des valeurs des caractères. On rappelle que si (V, ρ) est une représentation de G, on a une structure de $K[G]$-module sur V donnée par

$$\Big(\sum_{g \in G} \lambda_g e_g\Big) * v = \sum_{g \in G} \lambda_g (g \cdot v).$$

On a le résultat important suivant.

5.1. Théorème. *Soit G un groupe fini, soit K un corps algébriquement clos de caractéristique nulle et soit une représentation irréductible (V, ρ), de caractère χ. Alors, pour toute classe de conjugaison C de G, il existe un unique élément $\omega_\chi(C) \in K$ tel que*

$$\sum_{g \in C} \rho(g) = \omega_\chi(C) \mathrm{Id}_V.$$

De plus :

(1) *on a* $\omega_\chi(C) = \dfrac{|C|\chi(g_C)}{\chi(1_G)}$, *où* $g_C \in C$ *est un élément arbitraire de* C ;

(2) *si* $C_1, \ldots, C_s$ *sont les classes de conjugaison de* G, *et si* $g_i \in C_i$, *pour tous* $i, j \in [\![1, s]\!]$, *on a*

$$\omega_\chi(C_i)\omega_\chi(C_j) = \sum_{k=1}^{s} \alpha_{ij}^{(k)} \omega_\chi(C_k),$$

avec $\alpha_{ij}^{(k)} = |\{(g, h) \in C_i \times C_j \mid gh = g_k\}|$;

(3) *l'élément* $\omega_\chi(C)$ *est entier sur* $\mathbb{Z}$.

Démonstration. Avant de commencer, notons que pour tous $\lambda_1, \lambda_2 \in K$, on a

$$\lambda_1 \mathrm{Id}_V = \lambda_2 \mathrm{Id}_V \quad \Longrightarrow \quad \lambda_1 = \lambda_2.$$

En effet, soit $v \in V \setminus \{0\}$. On a alors $(\lambda_1 - \lambda_2)v = 0$. Si $\lambda_1 \neq \lambda_2$, en multipliant par $(\lambda_1 - \lambda_2)^{-1}$, on obtient la contradiction $v = 0$. Par conséquent, on a bien $\lambda_1 = \lambda_2$. Autrement dit, l'élément $\lambda \in K$ est unique. Passons maintenant à la démonstration proprement dite.

Posons $u_C = \sum_{g \in C} \rho(g)$. Alors, pour tout $v \in V$, on a

$$u_C(v) = \sum_{g \in C} g \cdot v = e_C * v.$$

Comme $e_C \in Z(K[G])$ par le lemme 1.4, pour tout $g \in G$, et pour tout vecteur $v \in V$, on a[5]

$$u_C(g \cdot v) = e_C * (e_g * v) = (e_C * e_g) * v = (e_g * e_C) * v = e_g * (e_C * v) = g \cdot u_C(v).$$

Ainsi, u_C est un morphisme de représentations.

Par le lemme de Schur, u_C est une homothétie, et il existe ainsi un unique élément $\omega_\chi(C) \in K$ tel que

$$\sum_{g \in C} \rho(g) = \omega_\chi(C) \mathrm{Id}_V.$$

5. Cette égalité peut bien sûr se démontrer par calcul direct.

En prenant la trace des deux côtés de l'égalité, on obtient

$$\sum_{g\in C}\chi(g)=\dim_K(V)\omega_\chi(C).$$

Comme χ est constant sur C, en choisissant un élément $g_C\in C$, on obtient finalement

$$|C|\chi(g_C)=\dim_K(V)\omega_\chi(C),$$

d'où (1) puisque K est de caractéristique nulle, en tenant compte du fait que $\chi(1_G)=\dim_K(V)$.

Montrons (2). On a

$$u_{C_i}\circ u_{C_j}=\omega_\chi(C_i)\omega_{C_j}(C)\mathrm{Id}_V.$$

Par ailleurs, pour tout $v\in V$, on a

$$(u_{C_i}\circ u_{C_j})(v)=u_{C_i}\big(u_{C_j}(v)\big)=u_{C_i}(e_{C_j}*v)=e_{C_i}*(e_{C_j}*v),$$

soit encore

$$(u_{C_i}\circ u_{C_j})(v)=(e_{C_i}*e_{C_j})*v,\qquad \text{pour tout } v\in V.$$

Or, un simple calcul montre que

$$e_{C_i}*e_{C_j}=\sum_{k=1}^{s}\alpha_{ij}^{(k)}e_{C_k},$$

avec $\alpha_{ij}^{(k)}=|\{(g,h)\in C_i\times C_j\mid gh=g_k\}|$. On en déduit alors que

$$(u_{C_i}\circ u_{C_j})(v)=\omega_\chi(C_i)\omega_\chi(C_j)v=\Big(\sum_{k=1}^{s}\alpha_{ij}^{(k)}\omega_\chi(C_k)\Big)v,\qquad \text{pour tout } v\in V.$$

D'après la remarque préliminaire, on obtient (2).

Si $i\in[\![1,s]\!]$, posons $M_i=(\alpha_{ij}^{(k)})_{j,k}\in \mathrm{M}_s(K)$. Le point (2) se récrit alors

$$\omega_\chi(C_i)\begin{pmatrix}\omega_\chi(C_1)\\ \vdots\\ \omega_\chi(C_s)\end{pmatrix}=M_i\begin{pmatrix}\omega_\chi(C_1)\\ \vdots\\ \omega_\chi(C_s)\end{pmatrix}.$$

Ainsi, $\omega_\chi(C_i)$ est une valeur propre de M_i. Comme M_i est à coefficients entiers, le polynôme caractéristique de M_i est un polynôme unitaire à coefficients dans $\mathbb{Z}$, dont $\omega_\chi(C_i)$ est une racine. Par conséquent, $\omega_\chi(C_i)$ est entier sur $\mathbb{Z}$. Cela achève la démonstration. □

Notons maintenant que si $z\in Z(G)$, alors $\{z\}$ est une classe de conjugaison de G. On écrira alors $\omega_\chi(z)$ plutôt que $\omega_\chi(\{z\})$. On a alors le corollaire suivant.

5.2. Corollaire. *Soit G un groupe fini, soit K un corps algébriquement clos de caractéristique nulle et soit une représentation irréductible (V, ρ) de G, de caractère χ. Alors :*

(1) *pour tout $z \in Z(G)$, $\omega_\chi(z) \in K^\times$ et on a $\rho(z) = \omega_\chi(z)\mathrm{Id}_V$;*

(2) *pour tout $z \in Z(G)$ et tout $g \in G$, on a $\chi(zg) = \omega_\chi(z)\chi(g)$;*

(3) *l'application*

$$\begin{aligned} Z(G) &\longrightarrow K^\times \\ z &\longmapsto \omega_\chi(z) \end{aligned}$$

est un morphisme de groupes.

Démonstration. Soit $z \in Z(G)$. L'égalité $\rho(z) = \omega_\chi(z)\mathrm{Id}_V$ découle du théorème 5.1, puisque $\{z\}$ est une classe de conjugaison de G. Qui plus est, puisque $\rho(z) \in \mathrm{GL}(V)$, on a nécessairement $\omega_\chi(z) \in K^\times$.

Si $g \in G$, on a alors

$$\rho(zg) = \rho(z) \circ \rho(g) = \omega_\chi(z)\rho(g),$$

d'où $\chi(zg) = \omega_\chi(z)\chi(g)$ en prenant la trace.

Enfin, si z_1, $z_2 \in Z(G)$, le point (1) montre alors que

$$\rho(z_1 z_2) = \rho(z_1) \circ \rho(z_2) = \omega_\chi(z_1)\omega_\chi(z_2)\mathrm{Id}_V,$$

d'où l'égalité $\omega_\chi(z_1 z_2) = \omega_\chi(z_1)\omega_\chi(z_2)$ par définition de ω_χ. Cela achève la démonstration. □

Nous allons maintenant donner quelques applications du théorème précédent. On commence par un théorème dû à Burnside.

5.3. Théorème. *On suppose que $K = \mathbb{C}$.*

Soit G un groupe fini, et soit (V, ρ) une représentation irréductible de G, de caractère χ. On suppose que C est une classe de conjugaison de G telle que $\mathrm{pgcd}\big(|C|, \chi(1_G)\big) = 1$.

Alors, pour tout $g \in C$, soit $\chi(g) = 0$, soit $\rho(g)$ est une homothétie.

Démonstration. Par hypothèse, il existe u, $v \in \mathbb{Z}$ tels que $u|C| + v\chi(1_G) = 1$. Soit $g \in C$. En multipliant par $\dfrac{\chi(g)}{\chi(1_G)}$, on obtient

$$u\omega_\chi(C) + v\chi(g) = \frac{\chi(g)}{\chi(1_G)}.$$

Puisque $\omega_\chi(C)$ et $\chi(g)$ sont entiers sur $\mathbb{Z}$ d'après le théorème 5.1 (3) et le lemme 1.7 (1), on obtient finalement que $\dfrac{\chi(g)}{\chi(1_G)}$ est entier sur $\mathbb{Z}$ par le corollaire XXVI-4.6. Soit $e = \exp(G)$.

Comme $\chi(g)$ est une somme de n racines e-ièmes de l'unité, il en est de même du nombre complexe $\sigma(\chi(g))$ pour tout $\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta_e)/\mathbb{Q})$, d'après le lemme XXV-4.11. En particulier, on a

$$\Big|\frac{\sigma(\chi(g))}{\chi(1_G)}\Big| \leqslant 1.$$

De plus, l'élément

$$c = \prod_{\sigma} \frac{\sigma(\chi(g))}{\chi(1_G)}$$

est fixé par tout élément de $\mathrm{Gal}(\mathbb{Q}(\zeta_e)/\mathbb{Q})$. C'est donc un rationnel par le lemme XXVIII-2.10. Or, puisque $\dfrac{\chi(g)}{\chi(1_G)}$ est entier sur $\mathbb{Z}$,

$$\sigma\Big(\frac{\chi(g)}{\chi(1_G)}\Big) = \frac{\sigma(\chi(g))}{\chi(1_G)}$$

est aussi entier sur $\mathbb{Z}$. Par le corollaire XXVI-4.6, c est entier sur $\mathbb{Z}$. Mais, $\mathbb{Z}$ étant intégralement clos (car il est factoriel ; cf. exemple XXVI-4.9 (1)), on obtient $c \in \mathbb{Z}$.

Si $\Big|\dfrac{\chi(g)}{\chi(1_G)}\Big| < 1$, on obtient $|c| < 1$ d'après la majoration ci-dessus. Comme on a $c \in \mathbb{Z}$, on en déduit que $c = 0$. Ainsi, il existe $\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta_e)/\mathbb{Q})$ tel que $\sigma(\chi(g)) = 0$, d'où $\chi(g) = 0$ puisque σ est injective.

Si $\Big|\dfrac{\chi(g)}{\chi(1_G)}\Big| = 1$, on obtient $\rho(g) = z\mathrm{Id}_V$, avec $z \in \mu_e(\mathbb{C})$, et ce d'après le lemme 1.7 (2). Cela achève la démonstration. □

On peut maintenant démontrer le résultat suivant.

5.4. Théorème. *Soit G un groupe fini simple non abélien. Alors, la seule classe de conjugaison de G dont le cardinal est une puissance d'un nombre premier est $\{1_G\}$.*

Démonstration. Supposons qu'il existe une classe de conjugaison C telle que $|C| = p^r$, où p est premier et $r \geqslant 1$. Soit $\rho : G \longrightarrow \mathrm{GL}(V)$ une représentation irréductible sur $\mathbb{C}$ non triviale[6], de caractère χ. Notons que l'on a $\chi \neq \chi_0$, puisque ρ est non triviale (cf. remarque 1.6).

Puisque ρ est non triviale et que le groupe G est simple, on a $\mathrm{Ker}(\rho) = \{1_G\}$. Soit $g \in C$. Si $\rho(g)$ est une homothétie, alors $\rho(g)$ commute avec tout élément de $\mathrm{im}(\rho)$. Comme ρ est injective, on en déduit que g commute avec

6. Une telle représentation existe par le corollaire XXXIV-5.2, puisque G est non abélien.

tous les éléments de G, i.e. $g \in Z(G)$. Comme G est simple et non abélien, on a $Z(G) = \{1_G\}$, et donc $g = 1_G$.

Deux cas se présentent alors. Si $\operatorname{pgcd}\big(|C|, \chi(1_G)\big) \neq 1$, on a $p \mid \chi(1_G)$. Si $\operatorname{pgcd}\big(|C|, \chi(1_G)\big) = 1$, par le théorème 5.3, on déduit de ce qui précède que $\chi(g) = 0$, pour tout $g \in C \setminus \{1_G\}$.

Soit X l'ensemble des caractères irréductibles $\chi \neq \chi_0$ tels que $p \mid \chi(1_G)$.

Fixons $g \in C \setminus \{1_G\}$. Par le corollaire 2.15 par exemple, on a

$$0 = \sum_{\chi \in \mathrm{Irr}(G)} \chi(g)\chi(1_G).$$

D'après ce qui précède, on obtient

$$0 = \chi_0(g)\chi_0(1_G) + \sum_{\chi \in X} \chi(g)\chi(1_G) = 1 + \sum_{\chi \in X} \chi(g)\chi(1_G).$$

On obtient alors

$$\frac{1}{p} = -\sum_{\chi \in X} \frac{\chi(1_G)}{p}\chi(g).$$

Comme $\chi(g)$ est entier sur $\mathbb{Z}$ et $\dfrac{\chi(1_G)}{p}$ est un entier pour tout $\chi \in X$, on en conclut que $\dfrac{1}{p}$ est entier sur $\mathbb{Z}$, ce qui est absurde, puisque les seuls éléments de $\mathbb{Q}$ entiers sur $\mathbb{Z}$ sont les éléments de $\mathbb{Z}$. Cela achève la démonstration. □

Nous sommes maintenant en mesure de démontrer un célèbre résultat de théorie des groupes, dû à Burnside.

5.5. Théorème (Théorème $p^a q^b$). *Soit G un groupe fini d'ordre $p^a q^b$, où p, q sont premiers et a, $b \geqslant 0$. Alors, G est résoluble.*

Démonstration. Nous allons procéder par récurrence sur $|G|$. Pour tout entier $n \geqslant 1$, soit (H_n) la propriété suivante :

(H_n) tout groupe fini G, tel que $|G| = p^a q^b \leqslant n$, est résoluble.

Si $n = 1$, G est trivial, donc résoluble.

Supposons que (H_n) soit vraie pour un certain entier $n \geqslant 1$, et soit G un groupe d'ordre $p^a q^b \leqslant n + 1$. On peut toujours supposer que $|G| = n + 1$ (sans quoi, on utilise l'hypothèse (H_n)). Puisqu'un p-groupe est résoluble par le théorème IX-3.23, on peut toujours supposer que $p \neq q$ et a, $b \geqslant 1$. De même, on peut supposer que G n'est pas abélien.

Soit S un p-Sylow de G. Puisque $a \geqslant 1$, $Z(S)$ est non trivial par la proposition VII-2.7. Soit $g_0 \in Z(S) \setminus \{1_G\}$. On note C sa classe de conjugaison.

Comme g_0 commute avec tout élément de S, on obtient

$$S \subset C_G(g_0) = \{g \in G \mid g g_0 g^{-1} = g_0\}.$$

Remarquons maintenant que $C_G(g_0)$ est le stabilisateur de g_0 sous l'action de G sur lui-même par conjugaison. Comme l'orbite de g_0 sous cette action est C, on a l'égalité

$$|C| = [G : C_G(g_0)].$$

Comme $S \subset C_G(g_0)$, on a $[G : C_G(g_0)] \mid [G : S]$, ce qui se récrit $|C| \mid q^b$. Notons que C n'est pas réduite à $\{1_G\}$, puisque $g_0 \neq 1_G$. Comme $|C|$ est une puissance non triviale d'un nombre premier et que G est non abélien, le théorème 5.4 montre que G n'est pas simple. Il existe donc un sous-groupe H distingué de G, distinct de $\{1_G\}$ et G. Mais alors, les ordres de H et G/H sont des diviseurs stricts de $|G|$, et vérifient ainsi (H_n). Par hypothèse de récurrence, H et G/H sont résolubles. Mais alors, G est résoluble par la proposition IX-3.21 (3). Cela achève la récurrence, ainsi que la démonstration. □

Nous achevons cette section en démontrant un théorème d'annulation pour les caractères.

5.6. Proposition. *Soit G un groupe fini. On suppose que $K = \mathbb{C}$. Alors, le caractère d'une représentation irréductible de G de degré $n \geqslant 2$ s'annule au moins une fois.*

Démonstration. Soit χ un caractère de G. Pour tout $g \in G$, on note

$$X(g) = \big\{g^k \mid k \in [\![1, o(g)]\!],\ \operatorname{pgcd}\big(k, o(g)\big) = 1\big\}.$$

Commençons par démontrer que $\displaystyle\prod_{g' \in X(g)} \chi(g') \in \mathbb{Z}$.

Soit $g \in G$, et soit $d = o(g)$. Soit $k \in [\![1, d]\!]$, $\operatorname{pgcd}(k, d) = 1$. D'après le lemme XXV-4.11, il existe un unique automorphisme $\sigma_k \in \operatorname{Gal}\big(\mathbb{Q}(\zeta_d)/\mathbb{Q}\big)$ tel que $\sigma_k(\zeta_d) = \zeta_d^k$. Comme toute racine d-ième de l'unité est une puissance de ζ_d, on a aussi

$$\sigma_k(\zeta) = \zeta^k, \quad \text{pour tout } \zeta \in \mu_d(\mathbb{C}).$$

Or, l'endomorphisme $\rho(g)$ est diagonalisable, et ses valeurs propres $z_1, \ldots, z_n$ sont des racines d-ièmes de l'unité par le lemme 1.7 (1). On en déduit alors que les valeurs propres de $\rho(g^k) = \rho(g)^k$ sont $z_1^k, \ldots, z_n^k$. Par conséquent, on obtient

$$\chi(g^k) = z_1^k + \cdots + z_n^k = \sigma_k(z_1) + \cdots + \sigma_k(z_n) = \sigma_k\big(\chi(g)\big).$$

Comme les éléments du groupe de Galois $\operatorname{Gal}\big(\mathbb{Q}(\zeta_d)/\mathbb{Q}\big)$ sont exactement

les automorphismes σ_k, avec $k \in [\![1, d]\!]$ tel que $\operatorname{pgcd}(k, d) = 1$, on a

$$\prod_{g' \in X(g)} \chi(g') = \prod_{\sigma \in \mathrm{Gal}(\mathbb{Q}(\zeta_d)/\mathbb{Q})} \sigma\big(\chi(g)\big).$$

En particulier, $\prod_{g' \in X(g)} \chi(g')$ est fixé par tous les éléments de $\mathrm{Gal}\big(\mathbb{Q}(\zeta_d)/\mathbb{Q}\big)$. C'est donc un rationnel par le lemme XXVIII-2.10. Il est aussi un élément entier sur $\mathbb{Z}$ par le lemme 1.7 (1) (et le corollaire XXVI-4.6) ; c'est donc un élément de $\mathbb{Z}$.

Supposons maintenant que χ soit un caractère irréductible ne s'annulant pas. Alors, $\prod_{g' \in X(g)} \chi(g')$ est un entier non nul, et ainsi

$$\Big(\prod_{g' \in X(g)} |\chi(g')| \Big)^2 \geqslant 1.$$

L'inégalité arithmético-géométrique donne alors

$$\frac{1}{|X(g)|} \sum_{g' \in X(g)} |\chi(g')|^2 \geqslant \prod_{g' \in X(g)} |\chi(g')|^{\frac{2}{|X(g)|}} \geqslant 1,$$

d'où finalement

$$\sum_{g' \in X(g)} |\chi(g')|^2 \geqslant |X(g)|.$$

Remarquons maintenant que G est la réunion disjointe[7] d'ensembles de la forme $X(g)$, avec $g \in G$. Notons aussi que $X(1_G) = \{1_G\}$. On en déduit alors que

$$\sum_{g \in G \setminus \{1_G\}} |\chi(g)|^2 \geqslant \sum_{g \in G \setminus \{1_G\}} |X(g)|,$$

et ainsi

$$\sum_{g \in G} |\chi(g)|^2 \geqslant \sum_{g \in G} |X(g)| - 1 + |\chi(1_G)|^2.$$

Par conséquent, on a

$$\sum_{g \in G} |\chi(g)|^2 \geqslant |G| - 1 + |\chi(1_G)|^2.$$

Comme χ est irréductible, on a

$$\langle \chi, \chi \rangle_G = 1$$

d'après la proposition 2.5 (1), par exemple.

7. En effet, il suffit de choisir un élément $g_1 \in G$, puis un élément $g_2 \in G \setminus X(g_1)$, puis un élément $g_3 \in G \setminus (X(g_1) \cup X(g_2))$, etc., et cela jusqu'à épuiser tous les éléments de G.

En tenant compte du lemme 3.1, cela se récrit

$$\sum_{g\in G} |\chi(g)|^2 = |G|.$$

Mais alors, on obtient

$$|G| \geqslant |G| - 1 + |\chi(1_G)^2|.$$

Comme $\chi(1_G)$ est un entier naturel, égal au degré d'une représentation irréductible associée à χ, on obtient $\chi(1_G) = 1$, d'où le résultat voulu. □

6. Degré des représentations irréductibles

Nous allons maintenant énoncer et démontrer des résultats permettant d'estimer les degrés des représentations irréductibles d'un groupe fini G. Certains de ces résultats seront obtenus grâce aux propriétés d'intégralité établies dans la section précédente.

On commence par le résultat le plus classique.

6.1. Théorème. *Soit G un groupe fini et soit K un corps algébriquement clos de caractéristique nulle. Alors, le degré d'une représentation irréductible divise $|G|$.*

Démonstration. Soient $C_1, \ldots, C_s$ les classes de conjugaison de G. Pour tout $i \in [\![1, s]\!]$, choisissons un élément $g_i \in C_i$. Soit V une représentation irréductible de G, de caractère χ. Alors, $\langle \chi, \chi \rangle_G = 1$,

$$|G| = \sum_{g\in G} \chi(g)\chi(g^{-1}) = \sum_{i=1}^{s} |C_i|\chi(g_i)\chi(g_i^{-1}),$$

puisque si un élément g est conjugué à g_i, g^{-1} est conjugué à g_i^{-1}. Mais alors, on obtient

$$\frac{|G|}{\chi(1_G)} = \sum_{i=1}^{s} \omega_\chi(C_i)\chi(g_i^{-1}).$$

Comme $\omega_\chi(C_i)$ et $\chi(g_i^{-1})$ sont entiers sur $\mathbb{Z}$ par le théorème 5.1 (3) et le lemme 1.7 (1), on en déduit que le rationnel $\dfrac{|G|}{\chi(1_G)}$ est entier sur $\mathbb{Z}$. C'est donc un entier, ce qui revient à dire que $\chi(1_G)$ divise $|G|$. Comme on a $\chi(1_G) = \dim_K(V)$, cela achève la démonstration. □

6.2. Remarque. On peut démontrer que ce résultat est vrai plus généralement si K est algébriquement clos et $\mathrm{car}(K) \nmid |G|$, mais la démonstration est nettement plus délicate et utilise des outils beaucoup plus sophistiqués[8]. En revanche, cela devient faux si $\mathrm{car}(K) \mid |G|$ (cf. exercice 14).

8. C'est, par exemple, une conséquence immédiate de la proposition 43 (ii) de [45].

Le résultat suivant permet également de réduire les possibilités.

6.3. Proposition. *Soit G un groupe fini, et soit K un corps quelconque tel que* $\mathrm{car}(K) \nmid |G|$. *Si H est un sous-groupe de G, on note m_H le degré maximum d'une représentation irréductible de H. Alors, pour tout sous-groupe de G, on a*

$$m_G \leqslant [G : H] m_H.$$

En particulier, si K est algébriquement clos et de caractéristique nulle, pour tout sous-groupe abélien A de G, le degré d'une représentation irréductible de G est inférieur ou égal à $[G : A]$.

Démonstration. Soit H un sous-groupe de G et soit (V, ρ) une représentation irréductible de G. Par restriction de l'action de G à H, V est aussi une représentation de H. Soit W une sous-représentation irréductible de H intervenant dans la décomposition de V (vue comme représentation de H). Alors, $\sum_{g \in G} \rho(g)(W)$ est un sous-espace de V non nul, stable par G. Par irréductibilité de V, on en déduit

$$V = \sum_{g \in G} \rho(g)(W).$$

Remarquons maintenant que si $g \in G$ et $h \in H$, on a

$$\rho(gh)(W) = \rho(g)\big(\rho(h)(W)\big) = \rho(g)(W).$$

Si $g_1, \ldots, g_r \in G$ forment un système complet de représentants de G/H, on en déduit l'égalité

$$V = \sum_{i=1}^{r} \rho(g_i)(W).$$

Mais alors, on obtient

$$\dim_K(V) \leqslant \sum_{i=1}^{r} \dim_K \big(\rho(g_i)(W)\big) \leqslant r \dim_K(W) \leqslant r m_H,$$

d'où la première partie en prenant une représentation irréductible V de dimension maximale. La seconde partie découle alors du fait que $m_A = 1$ lorsque K est algébriquement clos de caractéristique nulle, par la proposition 2.18. □

Notre prochain objectif est d'améliorer le théorème 6.1. Nous commençons par un lemme.

6.4. Lemme. *Soit G un groupe fini et soit K un corps algébriquement clos de caractéristique nulle. Enfin, soit $\chi \in \mathrm{Irr}(G)$ un caractère irréductible*

couplée au fait bien connu des spécialistes que tout corps de caractéristique p est isomorphe au corps résiduel d'un corps valué complet de caractéristique nulle.

de G. On suppose que $g \in G$ est conjugué à zg, pour un certain $z \in Z(G)$. Alors, soit $\chi(g) = 0$, soit $\chi(z) = \chi(1_G)$.

Démonstration. Gardons les notations et les hypothèses de l'énoncé. D'après le corollaire 5.2, on a

$$\chi(zg) = \omega_\chi(z)\chi(g) = \frac{\chi(z)}{\chi(1_G)}\chi(g).$$

Comme zg et g sont conjugués, on a aussi $\chi(zg) = \chi(g)$, d'où finalement

$$\Big(1 - \frac{\chi(z)}{\chi(1_G)}\Big)\chi(g) = 0.$$

Le résultat en découle immédiatement. □

Le théorème suivant améliore nettement l'énoncé du théorème 6.1.

6.5. Théorème. *Soit G un groupe fini et soit K un corps algébriquement clos de caractéristique nulle. Alors, le degré d'une représentation irréductible divise $[G : Z(G)]$.*

Démonstration. Soit (V, ρ) une représentation irréductible de G, de caractère χ.

Supposons tout d'abord que ρ soit une représentation fidèle. Dans ce cas, le lemme 1.7 montre que

$$\{g \in G \mid \chi(g) = \chi(1_G)\} = \{1_G\}.$$

Remarquons maintenant que si C est la classe de conjugaison de $g_0 \in G$, et si $z \in Z(G)$, alors

$$zC = \{zc \mid c \in C\} = \{z(gg_0g^{-1}) \mid g \in G\} = \{g(zg_0)g^{-1} \mid g \in G\},$$

et zC est alors la classe de conjugaison de zg_0. Ainsi, $Z(G)$ agit sur l'ensemble $\mathscr{C}$ des classes de conjugaison de G par multiplication à gauche.

Soit $C \in \mathscr{C}$ une classe de conjugaison de G, et soit $g \in C$. Si $\mathrm{Stab}_{Z(G)}(C)$ est non trivial, il existe $z \in Z(G) \setminus \{1_G\}$ tel que $zC = C$. Cela signifie que g et zg sont conjugués, et comme $z \neq 1_G$, on a $\chi(z) \neq \chi(1_G)$ d'après ce qui précède. D'après le lemme 6.4, on a alors $\chi(g) = 0$.

Pour toute classe de conjugaison C, choisissons un élément $g_C \in C$. La proposition 2.5 (1) montre alors que

$$|G| = \sum_{g \in G} \chi(g)\chi(g^{-1}) = \sum_{C \in \mathscr{C}} |C|\chi(g_C)\chi(g_C^{-1}) = \sum_{C \in \mathscr{C}'} |C|\chi(g_C)\chi(g_C^{-1}),$$

où $\mathscr{C}'$ est l'ensemble des classes de conjugaison de G telles que $\mathrm{Stab}_{Z(G)}(C)$ est trivial.

Notons que l'action de $Z(G)$ sur $\mathscr{C}$ se restreint en une action sur $\mathscr{C}'$. En effet, si $C \in \mathscr{C}'$ et $z \in Z(G)$, pour tout $z' \in Z(G)$, on a

$$z'(zC) = zC \iff (z^{-1}z'z)C = C \iff z^{-1}z'z = 1_G \iff z' = 1_G.$$

Soient alors $C_1, \ldots, C_t \in \mathscr{C}'$ un système complet de représentants des orbites de $\mathscr{C}'$. Autrement dit, les éléments de $\mathscr{C}'$ sont exactement les classes de conjugaison

$$zC_i, \quad z \in Z(G),\ i \in [\![1, t]\!],$$

toutes ces classes étant deux à deux distinctes.

Posant $g_i = g_{C_i}$ pour alléger les notations, on a donc

$$|G| = \sum_{i=1}^{t} \sum_{z \in Z(G)} |zC_i| \chi(zg_i) \chi\big((zg_i)^{-1}\big).$$

Soit $z \in Z(G)$ et soit $i \in [\![1, t]\!]$. D'après le corollaire 5.2, on a

$$\begin{aligned} \chi(zg_i)\chi\big((zg_i)^{-1}\big) &= \omega_\chi(z)\omega_\chi(z^{-1})\chi(g_i)\chi(g_i^{-1}) \\ &= \omega_\chi(z)\omega_\chi(z)^{-1}\chi(g_i)\chi(g_i^{-1}) \\ &= \chi(g_i)\chi(g_i^{-1}). \end{aligned}$$

Par ailleurs, la translation à gauche par z étant bijective, on a $|zC_i| = |C_i|$. On en déduit donc que

$$|G| = |Z(G)| \sum_{i=1}^{t} |C_i| \chi(g_i)\chi(g_i^{-1}) = |Z(G)| \chi(1_G) \sum_{i=1}^{t} \omega_\chi(C_i)\chi(g_i^{-1}),$$

soit encore

$$\frac{[G : Z(G)]}{\chi(1_G)} = \sum_{i=1}^{t} \omega_\chi(C_i)\chi(g_i^{-1}).$$

Par le théorème 5.1 (3) et le lemme 1.7 (1), on en déduit que $\dfrac{[G : Z(G)]}{\chi(1_G)}$ est entier sur $\mathbb{Z}$. C'est donc un entier, ce qui revient à dire que $\chi(1_G)$ divise $[G{:}Z(G)]$. Comme $\chi(1_G){=}\dim_K(V)$, on a le résultat voulu lorsque ρ est fidèle.

Soit maintenant une représentation irréductible $\rho : G \longrightarrow \mathrm{GL}(V)$ arbitraire, de caractère χ. Soit $N = \mathrm{Ker}(\rho)$. Alors, la représentation

$$\overline{\rho} : G/N \longrightarrow \mathrm{GL}(V)$$

est fidèle, et irréductible d'après l'exemple XXXIV-4.13. Le cas précédent montre alors que l'on a $\dim_K(V) \mid [G/N : Z(G/N)]$.

Or, il est facile de voir que l'image du centre $Z(G)$ de G par la projection canonique $G \longrightarrow G/N$ est incluse dans $Z(G/N)$. On obtient alors un

morphisme de groupes surjectif

$$G/Z(G) \longrightarrow (G/N)/Z(G/N).$$

En appliquant le premier théorème d'isomorphisme, on obtient en particulier que $[G/N : Z(G/N)] \mid [G : Z(G)]$. Mais alors, $\dim_K(V) \mid [G : Z(G)]$. Cela achève la démonstration. □

On termine cette section en généralisant à la fois le théorème 6.5 et la proposition 6.3.

6.6. Théorème. *Soit G un groupe fini et soit K un corps algébriquement clos de caractéristique nulle. Alors, pour tout sous-groupe abélien A distingué dans G, le degré d'une représentation irréductible de G divise $[G : A]$.*

Démonstration. Avant de démontrer le résultat, nous aurons besoin de quelques résultats intermédiaires.

Soit V une représentation irréductible de G et soit A un sous-groupe distingué de G. En restreignant l'action de G, V est aussi une représentation linéaire de A, mais plus nécessairement irréductible. Soit $W \subset V$ une sous-représentation irréductible de A, alors pour tout $g \in G$, $g{\cdot}W$ est encore une sous-représentation irréductible de A.

En effet, si W' est un sous-espace A-stable de $g{\cdot}W$, alors pour tout $a \in A$, on a

$$a{\cdot}(g^{-1}{\cdot}W') = g^{-1}{\cdot}\big((gag^{-1}){\cdot}W'\big) = g^{-1}{\cdot}W',$$

la dernière égalité utilisant le fait que A est distingué dans G. Ainsi, $g^{-1}{\cdot}W'$ est un sous-espace A-stable de W, donc égal à 0 ou W par irréductibilité de W. Mais alors, $W' = 0$ ou $g{\cdot}W$, d'où l'irréductibilité de $g{\cdot}W$.

Notons également que si W_1 et W_2 sont des sous-représentations de V, et si $f : W_1 \longrightarrow W_2$ est un isomorphisme de représentations de A, il en est de même de

$$\begin{aligned} f' : g{\cdot}W_1 &\longrightarrow g{\cdot}W_2 \\ g{\cdot}w_1 &\longmapsto g{\cdot}f(w_1). \end{aligned}$$

En effet, cette application est bien définie puisque l'application

$$\begin{aligned} V &\longrightarrow V \\ v &\longmapsto g{\cdot}v \end{aligned}$$

induit une bijection de W_1 sur $g{\cdot}W_1$. De plus, f' est bijective, comme composée de deux bijections.

Enfin, pour tout $a \in A$, et pour tout $w_1 \in W_1$, on a

$$\begin{aligned} f'\big(a\cdot(g\cdot w_1)\big) &= f'\big(g\cdot(g^{-1}ag)\cdot w_1\big) \\ &= g\cdot f(g^{-1}ag\cdot w_1) \\ &= g\cdot\big((g^{-1}ag)\cdot f(w_1)\big) \\ &= a\cdot\big(g\cdot f(w_1)\big) \\ &= a\cdot f'(g\cdot w_1). \end{aligned}$$

En appliquant ce qui précède à g^{-1}, on obtient en fait que pour toutes sous-représentations irréductibles $W_1 \subset V$ et $W_2 \subset V$ de A, on a

$$W_1 \simeq W_2 \iff g\cdot W_1 \simeq g\cdot W_2.$$

Considérons maintenant la décomposition de V en composantes isotypiques

$$V = V^{(1)} \oplus \cdots \oplus V^{(r)}.$$

On a alors

$$V = g\cdot V = g\cdot V^{(1)} \oplus \cdots \oplus g\cdot V^{(r)}.$$

Soit W_i une représentation irréductible de A telle que $V^{(i)}$ soit somme directe de sous-représentations irréductibles toutes isomorphes à W_i. Le point précédent montre que $g\cdot V^{(i)}$ est une somme directe de sous-représentations irréductibles toutes isomorphes à $g\cdot W_i$. De plus, si $i \neq j$, alors $W_i \not\simeq W_j$ par définition des composantes isotypiques, et par conséquent $g\cdot W_i \not\simeq g\cdot W_j$.

Bref, pour tout $i \in [\![1, r]\!]$, et tout $g \in G$, $g\cdot V^{(i)}$ est encore une composante isotypique de V. De plus, G permute transitivement $V^{(1)}, \ldots, V^{(r)}$. En effet, dans le cas contraire, on obtiendrait une décomposition de V en deux sous-espaces non nuls G-stables, ce qui contredirait l'irréductibilité de V.

Posons maintenant

$$H = \{g \in G \mid g\cdot V^{(1)} = V^{(1)}\}.$$

Comme $V^{(1)}$ est une représentation de A, on a $A \subset H$. Soit $\{g_1, \ldots, g_t\}$ un système complet de représentants de G/H. D'après ce qui précède, toute composante isotypique est de la forme $g\cdot V^{(1)}$, avec $g \in G$. Par définition, on a $g = g_j h$, avec $h \in H$ et $j \in [\![1, t]\!]$. Mais alors, on a $g\cdot V^{(1)} = g_j\cdot V^{(1)}$ par définition de H. De plus, si $g_k\cdot V^{(1)} = g_\ell\cdot V^{(1)}$, alors $g_k g_\ell^{-1} \in H$. Par choix de $g_1, \ldots, g_t$, on obtient $g_k = g_\ell$. Ainsi, les composantes isotypiques de V sont exactement les sous-espaces $g_j\cdot V^{(1)}, j \in [\![1, t]\!]$. En particulier, on a les égalités $r = t = [G : H]$. On a donc

$$V = g_1\cdot V^{(1)} \oplus \cdots \oplus g_t\cdot V^{(1)}.$$

Notons enfin que, par définition de H, l'action de G sur V se restreint en une action de H sur $V^{(1)}$. De plus, $V^{(1)}$ est une représentation irréductible de H.

En effet, si W est un sous-espace H-stable de $V^{(1)}$, alors $g_1 \cdot W \oplus \cdots \oplus g_t \cdot W$ est un sous-espace G-stable de V. En particulier, si $V^{(1)}$ se décompose en somme directe interne de deux sous-espaces H-stables non nuls, cela induit une décomposition de V en somme directe interne de deux sous-espaces G-stables non nuls, ce qui n'est pas possible par irréductibilité de V. Remarquons au passage que l'on a l'égalité

$$\dim_K(V) = [G : H] \dim_K(V^{(1)}).$$

Nous allons maintenant pouvoir démontrer le résultat voulu. On va procéder par récurrence sur $|G|$. Pour tout $n \geqslant 1$, soit (H_n) la propriété suivante :

(H_n) pour tout groupe G tel que $|G| \leqslant n$, tout sous-groupe abélien A distingué dans G, et toute représentation irréductible V, on a

$$\dim_K(V) \mid [G : A].$$

La propriété (H_1) est trivialement vraie. Supposons que (H_n) soit vraie pour un entier $n \geqslant 1$, et montrons que (H_{n+1}) est vraie. Soit G un groupe tel que $|G| \leqslant n+1$. Soit A un sous-groupe abélien distingué dans G et soit V une représentation irréductible de G.

Gardons les notations précédentes. Supposons tout d'abord que $H \neq G$. On a donc $|H| \leqslant n$. Comme $A \subset H$, A est aussi un sous-groupe distingué de H. Puisque $V^{(1)}$ est une représentation irréductible de H, l'hypothèse de récurrence montre que $\dim_K(V^{(1)}) \mid [H : A]$. En multipliant par $[G : H]$, on en déduit que $\dim_K(V) \mid [G : A]$.

Supposons maintenant que l'égalité $G = H$ ait lieu, autrement dit que $t = 1$. Alors, $V = V^{(1)}$, et l'on a

$$V = W_1 \oplus \cdots \oplus W_k,$$

où chaque W_m est isomorphe à une même représentation irréductible W de A. Comme A est abélien, chaque représentation W_m est de degré 1 par la proposition 2.18. En particulier, $\rho(a)_{|W_m}$ est une homothétie de W_m pour tout $a \in A$. Comme $W_1, \ldots, W_m$ sont toutes isomorphes, on constate aisément que $\rho(a)_{|W_1}, \ldots, \rho(a)_{|W_k}$ sont des homothéties de même rapport. Par conséquent, $\rho(a)$ est une homothétie de V pour tout $a \in A$. En particulier, $\rho(A) \subset Z\big(\rho(G)\big)$.

Remarquons que V est une représentation irréductible[9] de $\rho(G)$ (où l'action de $\rho(G)$ sur V est induite par l'action naturelle de $\mathrm{GL}(V)$). D'après le théorème précédent, on a donc $\dim_K(V) \mid \big[\rho(G) : Z\big(\rho(G)\big)\big]$.

9. La vérification en est laissée au lecteur.

Or, puisque $\rho(A) \subset Z\big(\rho(G)\big)$, le morphisme surjectif $\rho : G \longrightarrow \rho(G)$ induit un morphisme surjectif $G/A \longrightarrow \rho(G)/Z\big(\rho(G)\big)$. On a donc

$$\big[\rho(G) : Z\big(\rho(G)\big)\big] \mid [G : A].$$

On en déduit alors que $\dim_K(V) \mid [G : A]$.

Bref, dans tous les cas, on a démontré que (H_{n+1}) est vraie. Cela achève la récurrence, ainsi que la démonstration [10]. □

7. Exercices

E1. Donner un exemple de représentation irréductible V d'un groupe G qui ne soit pas irréductible lorsqu'on la restreint à un sous-groupe H bien choisi.

E2. Soit K un corps infini de caractéristique différente de 2 et 3.

a. On suppose que K est algébriquement clos. Montrer que les représentations irréductibles de $\mathfrak{S}_3$ sont $\mathbf{1}$, ε et H_3.

b. On revient au cas général. Soit V une représentation de degré fini de $\mathfrak{S}_3$ sur $\overline{K}$. Déduire de la question précédente qu'il existe une représentation V' de degré fini de $\mathfrak{S}_3$ sur K telle que $V \simeq V' \otimes_K \overline{K}$.

c. Montrer que le résultat de la question **a.** est encore vrai si K n'est pas algébriquement clos.

E3. Soit $n \geqslant 2$. On suppose que $\mathrm{car}(K) \neq 2$. Montrer que les représentations irréductibles H_n et $H_n(\varepsilon)$ de $\mathfrak{S}_n$ sont isomorphes si, et seulement si, $n = 3$.

E4. Soit G un groupe fini. Si χ est le caractère d'une représentation irréductible de G sur $\mathbb{C}$, on note $\overline{\chi}$ l'application définie par

$$\overline{\chi}(g) = \overline{\chi(g)}, \quad \text{pour tout } g \in G.$$

a. Justifier que si χ est le caractère d'une représentation irréductible de G sur $\mathbb{C}$, il en est de même de $\overline{\chi}$.

Dans la suite, on dira que deux caractères irréductibles χ et χ' de G sont conjugués si $\chi' = \overline{\chi}$.

10. La démonstration présentée ici est une version très détaillée de celle proposée dans [45].

b. Soit V une représentation de G sur $\mathbb{R}$, de degré fini. Montrer que son caractère χ vérifie $\overline{\chi} = \chi$.

On note $\chi_1, \ldots, \chi_r$, $\chi'_1, \overline{\chi}'_1, \ldots, \chi'_t, \overline{\chi}'_t$ les caractères irréductibles de G sur $\mathbb{C}$, où $\overline{\chi}_i = \chi_i$ pour tout $i \in [\![1, r]\!]$, $\overline{\chi}'_j \neq \chi'_j$ pour tout $j \in [\![1, t]\!]$ et $\chi'_1, \ldots, \chi'_t$ sont deux à deux non conjugués.

c. Soit χ le caractère d'une représentation V de G sur $\mathbb{R}$. Montrer qu'il existe des entiers naturels $m_1, \ldots, m_r$, $m'_1, \ldots, m'_t$ tels que

$$\chi = \sum_{i=1}^{r} m_i \chi_i + \sum_{j=1}^{t} m'_j (\chi'_j + \overline{\chi}'_j).$$

Indication. Utiliser le lemme d'indépendance de Dedekind.

E5. Soient G_1, G_2 deux groupes et soit K un corps. Pour $i = 1, 2$, on se donne une représentation $\rho_i : G_i \longrightarrow \mathrm{GL}(V_i)$.

a. Montrer qu'il existe une unique représentation

$$\rho_1 \boxtimes \rho_2 : G_1 \times G_2 \longrightarrow \mathrm{GL}(V_1 \otimes_K V_2)$$

vérifiant

$$(\rho_1 \boxtimes \rho_2)\big((g_1, g_2)\big) = \rho_1(g_1) \otimes \rho_2(g_2), \quad \text{pour tout } (g_1, g_2) \in G_1 \times G_2.$$

b. Si χ_i est le caractère de ρ_i, montrer que le caractère $\chi_1 \boxtimes \chi_2$ de $\rho_1 \boxtimes \rho_2$ est donné par

$$(\chi_1 \boxtimes \chi_2)\big((g_1, g_2)\big) = \chi_1(g_1)\chi_2(g_2), \quad \text{pour tout } (g_1, g_2) \in G_1 \times G_2.$$

On suppose dorénavant que K est algébriquement clos de caractéristique nulle, et que G_1 et G_2 sont finis.

c. Montrer que $\rho_1 \boxtimes \rho_2$ est irréductible si, et seulement si, ρ_1 et ρ_2 sont irréductibles.

d. Montrer que toute représentation irréductible de $G_1 \times G_2$ est de la forme $\rho_1 \boxtimes \rho_2$, où ρ_i est une représentation irréductible de G_i.

Indication. On pourra calculer la somme des carrés des degrés des représentations $\rho_1 \boxtimes \rho_2$, lorsque ρ_i parcourt un système complet de représentations irréductibles de G_i.

E6. Soit $G = Q_8$.

a. Déterminer la table des caractères du groupe Q_8, et la comparer avec celle du groupe D_4.

b. Soit $\rho : Q_8 \longrightarrow \mathrm{GL}_2(\mathbb{C})$ la représentation irréductible de degré 2 donnée par l'inclusion.

Montrer que son caractère χ vérifie $\overline{\chi} = \chi$ (avec les notations de l'exercice 4 ci-dessus), mais qu'il n'existe pas de représentation

$$\rho' : Q_8 \longrightarrow \mathrm{GL}(V),$$

où V est un $\mathbb{R}$-espace vectoriel de dimension 2, telle que $\rho \simeq (\rho')_{\mathbb{C}}$.

E7. Soit $n \geqslant 2$. On écrit $n = 2\ell$ ou $2\ell+1$. Pour tout $k \in [\![1, \ell]\!]$, on considère la représentation de $\mathbb{Z}/n\mathbb{Z}$ définie par

$$\begin{array}{rccc} \rho_k : & \mathbb{Z}/n\mathbb{Z} & \longrightarrow & \mathrm{M}_2(\mathbb{R}) \\ & \overline{m} & \longmapsto & \begin{pmatrix} \cos\left(\frac{2mk\pi}{n}\right) & -\sin\left(\frac{2mk\pi}{n}\right) \\ \sin\left(\frac{2mk\pi}{n}\right) & \cos\left(\frac{2mk\pi}{n}\right) \end{pmatrix}. \end{array}$$

a. Montrer que ρ_k est irréductible, sauf si $n = 2\ell$ et $k = \ell$.

b. Soit χ_k le caractère associé à ρ_k. Montrer que $\langle \chi_k, \chi_k \rangle_G = 2$, sauf si $n = 2\ell$ et $k = \ell$, auquel cas $\langle \chi_k, \chi_k \rangle = 4$.

c. Décomposer $(\rho_k)_{\mathbb{C}}$ en somme directe de représentations irréductibles.

d. Utiliser l'exercice 4 pour montrer que les représentations irréductibles de $\mathbb{Z}/n\mathbb{Z}$ sur $\mathbb{R}$ sont :

(i) $\mathbf{1}$ et ρ_k, $k \in [\![1, \ell]\!]$ si $n = 2\ell + 1$;

(ii) $\mathbf{1}$, ρ_k, $k \in [\![1, \ell - 1]\!]$, et si $n = 2\ell$,

$$\begin{array}{rccc} \rho_{-1} : & \mathbb{Z}/n\mathbb{Z} & \longrightarrow & \mathbb{R}^\times \\ & \overline{m} & \longmapsto & (-1)^m. \end{array}$$

E8. Déterminer la table de caractères de $\mathfrak{S}_4$.

Indication. On pourra considérer l'action naturelle de $\mathfrak{S}_4$ sur $[\![1, 4]\!]$, et utiliser également l'exercice 3.

E9. Soit $G = \mathbb{Z}/3\mathbb{Z} \times_\rho \mathbb{Z}/4\mathbb{Z}$ l'unique produit semi-direct non trivial du groupe $\mathbb{Z}/3\mathbb{Z}$ par le groupe $\mathbb{Z}/4\mathbb{Z}$.

On rappelle que $G = \langle \sigma, \tau \rangle$, avec $o(\sigma) = 3$, $o(\tau) = 4$ tels que

$$\tau\sigma\tau^{-1} = \sigma^{-1}.$$

a. Montrer que $[G, G] = \langle \sigma \rangle$.

b. En déduire sans calculs que G possède quatre représentations irréductibles de degré 1 et deux représentations irréductibles de degré 2.

c. Montrer que G possède six classes de conjugaison, représentées respectivement par

$$g_1 = 1, \quad g_2 = \tau^2, \quad g_3 = \sigma, \quad g_4 = \sigma\tau^2, \quad g_5 = \tau, \quad g_6 = \tau^3.$$

d. Établir la table de caractères de G.

Indication. On pourra faire agir G sur l'ensemble de ses 2-Sylow.

E10. On suppose connue la table de caractères de $\mathfrak{S}_5$, et l'on se propose d'établir la table de caractères de $\mathfrak{A}_5$.

a. Vérifier que $\mathfrak{A}_5$ possède cinq classes de conjugaison, représentées par

$$g_1 = \mathrm{Id}, \quad g_2 = (1\ 2)(3\ 4), \quad g_3 = (1\ 2\ 3),$$
$$g_4 = (1\ 2\ 3\ 4\ 5), \quad g_5 = (1\ 3\ 5\ 4\ 2),$$

de cardinaux respectifs 1, 15, 20, 12 et 12.

b. Montrer que les restrictions des caractères χ_4 et χ_5 de $\mathfrak{S}_5$ sont encore des caractères irréductibles de $\mathfrak{A}_5$. On les note encore χ_4 et χ_5.

c. En déduire que les degrés des représentations irréductibles de $\mathfrak{A}_5$ sont égaux à 1, 3, 3, 4 et 5.

d. Montrer que la restriction ψ du caractère χ_6 de $\mathfrak{S}_5$ à $\mathfrak{A}_4$ vérifie

$$\langle \psi, \psi \rangle_{\mathfrak{A}_5} = 2, \quad \langle \psi, \chi_0 \rangle_{\mathfrak{A}_5} = 0, \quad \langle \psi, \chi_4 \rangle_{\mathfrak{A}_5} = 0, \quad \langle \psi, \chi_5 \rangle_{\mathfrak{A}_5} = 0.$$

En déduire que $\psi = \chi_3 + \chi'_3$, où χ_3 et χ'_3 sont deux caractères irréductibles, distincts de χ_0, χ_4 et χ_5. Montrer que χ_3 et χ'_3 sont associés à deux représentations irréductibles de degré 3 non isomorphes.

e. Calculer la norme des deuxième et troisième colonnes de la table de caractères, et en déduire que

$$\chi_3(g_2) = \chi'_3(g_2) = -1, \quad \chi_3(g_3) = \chi'_3(g_3) = 0.$$

f. Faire de même avec les quatrième et cinquième colonnes de la table de caractères, et compléter celle-ci.

E11. On se propose ici de déterminer toutes les représentations irréductibles du groupe D_n ($n \geqslant 3$) sur $\mathbb{C}$. On note r la rotation de centre O et d'angle $\frac{2\pi}{n}$, et s la réflexion orthogonale d'axe (Ox).

a. Vérifier que D_n possède un sous-groupe isomorphe à $(\mathbb{Z}/2\mathbb{Z})^2$ si n est pair, et à $\mathbb{Z}/2\mathbb{Z}$ si n est impair. Construire alors quatre représentations irréductibles de D_n de degré 1 si n est pair, et deux représentations irréductibles de degré 1 si n est impair. Calculer leurs caractères.

b. On pose $\zeta_n = e^{\frac{2i\pi}{n}}$. Pour tout $k \geqslant 0$, montrer qu'il existe une unique représentation $\rho_k : D_n \longrightarrow \mathrm{GL}_2(\mathbb{C})$ telle que

$$\rho_k(r) = \begin{pmatrix} \zeta_n^k & 0 \\ 0 & \zeta_n^{-k} \end{pmatrix} \quad \text{et} \quad \rho_k(s) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

c. On écrit $n = 2m$ ou $2m-1$. Montrer que pour tout $k \in [\![1, m-1]\!]$, les représentations $\rho_1, \ldots, \rho_m$ sont irréductibles et deux à deux non isomorphes.

d. Vérifier que l'on a obtenu ainsi toutes les représentations irréductibles de D_n à isomorphisme près.

e. La question précédente montre en particulier que les degrés des représentations irréductibles de D_n sont $\leqslant 2$. Comment aurait-on pu le prévoir a priori ?

f. Calculer $[D_n, D_n]$ et identifier $D_n/[D_n, D_n]$ (distinguer les cas selon la parité de n). Retrouver alors le fait que D_n possède exactement quatre représentations irréductibles de degré 1 si n est pair, et deux représentations irréductibles de degré 1 si n est impair.

E12. Soit $V = \mathbb{C}[D_4]$ la représentation régulière de D_4. Calculer les composantes isotypiques de V, et décomposer V en somme directe de sous-représentations irréductibles.

E13. Soit K un corps algébriquement clos et soit G un groupe fini. On se propose de démontrer que certains résultats sur les représentations irréductibles établis en caractéristique nulle sont encore vrais lorsque l'on suppose plus généralement que $\mathrm{car}(K) \nmid |G|$.

Soit $(V_1, \rho_1), \ldots, (V_s, \rho_s)$ un système complet de représentations irréductibles de G.

a. Montrer qu'il existe un unique morphisme de K-algèbres

$$\varphi : K[G] \longrightarrow \mathscr{L}(V_1) \times \cdots \times \mathscr{L}(V_s)$$

tel que

$$\varphi(e_g) = \big(\rho_1(g), \ldots, \rho_s(g)\big), \quad \text{pour tout } g \in G.$$

b. On considère l'application K-linéaire

$$\psi : \mathscr{L}(V_1) \times \cdots \times \mathscr{L}(V_s) \longrightarrow K[G]$$

définie par

$$\psi(u_1, \ldots, u_s) = \sum_{g \in G} \Big(\frac{1}{|G|} \sum_{i=1}^{s} n_i \mathrm{tr}\big(\rho_i(g^{-1}) u_i\big) \Big) e_g,$$

pour tout $u_i \in \mathscr{L}(V_i)$, où $n_i = \dim_K(V_i)$.

Montrer que φ et ψ sont inverses l'une de l'autre.

Indication. Utiliser le corollaire 2.3.

c. Montrer que l'on a un isomorphisme de K-algèbres

$$K[G] \simeq \mathrm{M}_{n_1}(K) \times \cdots \times \mathrm{M}_{n_s}(K).$$

d. En déduire que $|G| = n_1^2 + \cdots + n_s^2$ et que $s = |\mathrm{Conj}(G)|$.

Indication. Pour le second point, on pourra calculer la dimension du centre de $\mathrm{M}_{n_1}(K) \times \cdots \times \mathrm{M}_{n_s}(K)$.

E14. Soit p un nombre premier, et soit $G = \mathrm{SL}_2(\mathbb{F}_p)$.

Si $d \geqslant 0$, on note V_d l'ensemble $\overline{\mathbb{F}}_p[X,Y]_d$ des polynômes homogènes de degré d. C'est un $\overline{\mathbb{F}}_p$-espace vectoriel de dimension $d+1$, de base

$$(X^k Y^{d-k})_{k \in [\![0,d]\!]}.$$

On rappelle que le groupe linéaire $\mathrm{GL}_2(\overline{\mathbb{F}}_p)$ agit linéairement sur V_d (cf. exercice XXXIV-11). Par restriction, G agit également linéairement sur V_d, si bien que V_d est une représentation linéaire du groupe G, de degré $d+1$ sur $\overline{\mathbb{F}}_p$.

Explicitement, si $M = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in G$ et $P \in V_d$, on a

$$M \bullet P = P(dX - bY, -cX + aY).$$

On suppose dans toute la suite que $d < \dfrac{p-1}{2}$.

a. Montrer qu'il existe $\alpha \in \mathbb{F}_p^\times$ tel que les éléments α^{d-2k}, $k \in [\![0,d]\!]$, soient deux à deux distincts.

b. Soit W une sous-représentation de V_d.

Soit $P = \sum_{k \in S} \lambda_k X^k Y^{d-k} \in W$, où $S \subset [\![0,d]\!]$ et $\lambda_k \in \overline{K}^\times$ pour tout entier $k \in S$.

Montrer par récurrence sur $|S|$ que $X^k Y^{d-k} \in W$, pour tout $k \in S$.

Indication. Soit $\alpha \in \mathbb{F}_p^\times$ tel que les éléments α^{d-2k}, $k \in [\![0,d]\!]$ soient deux à deux distincts. Si $k \in S$, on pourra calculer

$$\alpha^{d-2k} P - \begin{pmatrix} \alpha & \\ & \alpha^{-1} \end{pmatrix} \bullet P,$$

et justifier que ce polynôme est un élément de W.

c. Soit W une sous-représentation de V_d non réduite à $\{0\}$.

(i) Justifier qu'il existe $i \in [\![0,d]\!]$ tel que $X^i Y^{d-i} \in W$.

(ii) Calculer $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \bullet X^i Y^{d-i}$. En utilisant **b.**, en déduire que $Y^d \in W$.

(iii) Calculer $\begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix} \bullet Y^d$, et en déduire que $X^k Y^{d-k} \in W$, pour tout entier $k \in [\![0,d]\!]$.

d. En déduire que, pour tout entier $d \geqslant 0$, V_d est une représentation irréductible de G de degré $d+1$ sur $\overline{\mathbb{F}}_p$, dont le degré ne divise pas nécessairement $|G|$. Donner un exemple numérique [11].

11. Merci à Vincent Beck pour m'avoir signalé ce contre-exemple.

Chapitre XXXVI

Introduction au problème de Noether

Dans tout ce qui suit, F désignera un corps arbitraire. Le point de départ de ce chapitre est le théorème de structure des polynômes symétriques (théorème XV-7.11), qui peut s'énoncer comme suit.

pour tout $P \in F[X_1, \ldots, X_n]$, les propriétés suivantes sont équivalentes :

(1) on a $P(X_{\tau(1)}, \ldots, X_{\tau(n)}) = P(X_1, \ldots, X_n)$ pour tout $\tau \in \mathfrak{S}_n$.

(2) il existe un unique $Q \in F[T_1, \ldots, T_n]$ tel que $P = Q(\sigma_1, \ldots, \sigma_n)$, où les polynômes $\sigma_1, \ldots, \sigma_n$ sont les polynômes symétriques élémentaires en $X_1, \ldots, X_n$.

On peut aussi démontrer un résultat similaire pour les fractions rationnelles, comme on le verra plus loin. Une question naturelle est de savoir si ces résultats sont encore vrais lorsque l'on remplace $\mathfrak{S}_n$ par un de ses sous-groupes.

Le but de ce chapitre est d'étudier succinctement cette question. Avant de l'aborder, nous allons en premier lieu faire quelques rappels sur les polynômes et fractions rationnelles.

1. Rappels et compléments sur les polynômes et les fractions rationnelles

Convention. Dans tout ce qui suit, toutes les F-algèbres seront implicitement supposées associatives et unitaires.

Le but de ce qui suit est de généraliser la construction du morphisme d'évaluation aux polynômes en plusieurs indéterminées et aux fractions rationnelles.

Notation. Si $\mathbf{m} = \begin{pmatrix} m_1 \\ \vdots \\ m_n \end{pmatrix} \in \mathbb{N}^n$, on pose $X^{\mathbf{m}} = X_1^{m_1} \cdots X_n^{m_n}$.

De même, si A est un anneau commutatif, pour tout $\mathbf{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \in A^n$, on pose $\mathbf{a}^{\mathbf{m}} = a_1^{m_1} \cdots a_n^{m_n}$.

On a alors la proposition suivante.

1.1. Proposition. *Soit $\mathscr{A}$ une F-algèbre commutative.*

Pour tout $\mathbf{a} = (a_1, \ldots, a_n) \in \mathscr{A}^n$, il existe un unique morphisme de F-algèbres $ev_{\mathbf{a}} : F[X_1, \ldots, X_n] \longrightarrow \mathscr{A}$, appelé morphisme d'évaluation en $\mathbf{a}$, tel que

$$ev_{\mathbf{a}}(X_i) = a_i, \quad \text{pour tout } i \in [\![1, n]\!].$$

Il est défini par

$$ev_{\mathbf{a}}\Big(\sum_{\mathbf{m}} \lambda_{\mathbf{m}} X^{\mathbf{m}}\Big) = \sum_{\mathbf{m}} \lambda_{\mathbf{m}} \mathbf{a}^{\mathbf{m}}.$$

De plus, tout morphisme de F-algèbres $f : F[X_1, \ldots, X_n] \longrightarrow \mathscr{A}$ est de la forme $ev_{\mathbf{a}}$, pour un unique $\mathbf{a} \in \mathscr{A}^n$.

Démonstration. Le fait que $ev_{\mathbf{a}}$ soit un morphisme de F-algèbres qui envoie X_i sur a_i est un calcul facile laissé au lecteur.

Si maintenant $f : F[X_1, \ldots, X_n] \longrightarrow \mathscr{A}$ est un morphisme de F-algèbres, on a nécessairement

$$ev_{\mathbf{a}}\Big(\sum_{\mathbf{m}} \lambda_{\mathbf{m}} X^{\mathbf{m}}\Big) = \sum_{\mathbf{m}} \lambda_{\mathbf{m}} f(X_1)^{m_1} \cdots f(X_n)^{m_n},$$

pour tout polynôme $P = \sum_{\mathbf{m}} \lambda_{\mathbf{m}} X^{\mathbf{m}} \in F[X_1, \ldots, X_n]$. Par conséquent, on obtient $f = ev_{\mathbf{a}}$, avec $a_i = f(X_i)$ pour tout $i \in [\![1, n]\!]$.

En particulier, si f est un morphisme de F-algèbres tel que $f(X_i) = a_i$, pour tout $i \in [\![1, n]\!]$, l'égalité précédente montre que $f = ev_{\mathbf{a}}$.

Enfin, si $f = ev_{\mathbf{a_1}} = ev_{\mathbf{a_2}}$, en calculant l'image de chaque X_i, on obtient immédiatement $\mathbf{a_1} = \mathbf{a_2}$. □

Notations. Si $P \in F[X_1, \ldots, X_n]$, on note $P(a_1, \ldots, a_n)$, voire $P(\mathbf{a})$, l'image de P par le morphisme $ev_{\mathbf{a}}$.

On note aussi $F[a_1, \ldots, a_n]$ l'image de $ev_{\mathbf{a}}$.

Lorsque $n = 1$ et $\mathbf{a}$ est un élément de $\mathscr{A}$, le morphisme $ev_{\mathbf{a}}$ est injectif si $\mathbf{a}$ est transcendant sur F. Lorsque $n \geqslant 1$, l'injectivité du morphisme d'évaluation conduit à la notion d'éléments algébriquement indépendants sur F.

1.2. Définition. Soit $\mathscr{A}$ une F-algèbre commutative. On dit que des éléments $a_1, \ldots, a_n \in \mathscr{A}$ sont *algébriquement indépendants sur* F si pour tout polynôme $P \in F[X_1, \ldots, X_n]$, on a

$$P(a_1, \ldots, a_n) = 0 \implies P = 0.$$

On s'intéresse maintenant au cas des fractions rationnelles, même si nous n'en aurons besoin que de façon anecdotique. Pour simplifier l'exposition, nous nous bornerons aux morphismes à valeurs dans un corps. On a alors la proposition suivante.

1.3. Proposition. *Soit L/F une extension de corps.*

Alors, pour tout n-uplet $\mathbf{a} = (a_1, \ldots, a_n) \in L^n$ tel que $a_1, \ldots, a_n$ sont des éléments algébriquement indépendants sur F, il existe un unique morphisme de F-algèbres $ev'_{\mathbf{a}} : F(X_1, \ldots, X_n) \longrightarrow L$, appelé morphisme d'évaluation en $\mathbf{a}$, tel que

$$ev'_{\mathbf{a}}(X_i) = a_i, \qquad \text{pour tout } i \in [\![1, n]\!].$$

Il est défini par

$$ev'_{\mathbf{a}}\Big(\frac{P}{Q}\Big) \;=\; P(\mathbf{a})Q(\mathbf{a})^{-1} \;\; \big(= P(a_1, \ldots, a_n)Q(a_1, \ldots, a_n)^{-1}\big).$$

De plus, tout morphisme de F-algèbres $f : F[X_1, \ldots, X_n] \longrightarrow L$ est de la forme $ev'_{\mathbf{a}}$, pour un unique $\mathbf{a} = (a_1, \ldots, a_n) \in L^n$, où $a_1, \ldots, a_n$ sont algébriquement indépendants sur F.

Démonstration. Soit $\mathbf{a} = (a_1, \ldots, a_n) \in L^n$, où $a_1, \ldots, a_n$ sont algébriquement indépendants sur F. Le morphisme $ev_{\mathbf{a}} : F[X_1, \ldots, X_n] \longrightarrow L$ est alors injectif. Ainsi, si $Q \in F[X_1, \ldots, X_n]$ est non nul, alors son évaluation en $\mathbf{a}$ est un élément non nul de L, donc inversible.

Si $f = \dfrac{P}{Q} \in F(X_1, \ldots, X_n)$, on vérifie[1] alors que la quantité $P(\mathbf{a})Q(\mathbf{a})^{-1}$ ne dépend pas de la représentation de f choisie, et que l'application

$$ev'_{\mathbf{a}} : F(X_1, \ldots, X_n) \longrightarrow L$$

est un morphisme de F-algèbres qui envoie X_i sur a_i.

Notons maintenant qu'un morphisme de F-algèbres

$$\rho : F(X_1, \ldots, X_n) \longrightarrow L$$

est non trivial. Comme $F(X_1, \ldots, X_n)$ est un corps, ce morphisme est nécessairement injectif. La restriction de ce morphisme à $F[X_1, \ldots, X_n]$ est donc un morphisme injectif $F[X_1, \ldots, X_n] \longrightarrow L$, donc de la forme $ev_{\mathbf{a}}$, où

1. On aurait pu aussi utiliser la proposition XV-3.18.

les coordonnées de $\mathbf{a}$ sont algébriquement indépendantes sur F. En particulier, un polynôme Q non nul est envoyé sur un élément non nul de L, donc inversible. Mais alors, on a

$$\rho\Big(\frac{P}{Q}\Big) = \rho(P)\rho(Q)^{-1} = P(\mathbf{a})Q(\mathbf{a})^{-1},$$

si bien que $\rho = ev'_{\mathbf{a}}$. L'unicité du n-uplet $\mathbf{a}$ se démontre comme dans le cas des polynômes. □

2. Polynômes et fractions rationnelles invariantes sous l'action d'un groupe fini

Comme déjà mentionné, le point de départ de tout ce qui suit est le théorème de structure des polynômes symétriques. Notons qu'un polynôme symétrique n'est rien d'autre qu'un point fixe de $F[X_1, \ldots, X_n]$ sous l'action de $\mathfrak{S}_n$ définie par

$$\begin{aligned} \mathfrak{S}_n \times F[X_1, \ldots, X_n] &\longrightarrow F[X_1, \ldots, X_n] \\ (\sigma, P) &\longmapsto \sigma{\cdot}P = P(X_{\sigma(1)}, \ldots, X_{\sigma(n)}). \end{aligned}$$

Par ailleurs, l'unicité de Q est de manière évidente équivalente à la propriété suivante : pour tout $Q \in F[T_1, \ldots, T_n]$, on a

$$Q(\sigma_1, \ldots, \sigma_n) = 0 \implies Q = 0.$$

Autrement dit, les polynômes $\sigma_1, \ldots, \sigma_n$ sont algébriquement indépendants sur F.

Le théorème précédent peut alors se réinterpréter comme suit : il existe des éléments $\sigma_1, \ldots, \sigma_n \in F[X_1, \ldots, X_n]^{\mathfrak{S}_n}$ vérifiant les deux conditions suivantes :

(1) on a $F[X_1, \ldots, X_n]^{\mathfrak{S}_n} = F[\sigma_1, \ldots, \sigma_n]$

(2) les polynômes $\sigma_1, \ldots, \sigma_n$ sont algébriquement indépendants sur F.

Une autre reformulation possible est la suivante : l'évaluation $T_i \rightsquigarrow \sigma_i$ induit un isomorphisme de F-algèbres

$$F[T_1, \ldots, T_n] \simeq F[X_1, \ldots, X_n]^{\mathfrak{S}_n}.$$

On a d'ailleurs une parfaite équivalence entre ces deux points de vue, comme le montre le lemme suivant.

2.1. Lemme. *Soit $\mathscr{A}$ une F-algèbre commutative et soit $r \geqslant 0$ un entier. Les propriétés suivantes sont équivalentes :*

(1) *il existe des éléments $a_1, \ldots, a_r \in \mathscr{A}$ algébriquement indépendants sur F tels que $\mathscr{A} = F[a_1, \ldots, a_r]$;*

(2) *il existe un isomorphisme de F-algèbres $F[T_1, \ldots, T_r] \simeq \mathscr{A}$.*

Démonstration. Supposons que (1) soit vérifiée, et soit $\mathbf{a} = (a_1, \ldots, a_r)$. Alors, le morphisme d'évaluation $ev_{\mathbf{a}}$ est injectif par hypothèse sur $\mathbf{a}$, et d'image $F[a_1, \ldots, a_r]$, d'où (2).

Inversement, supposons que (2) soit vérifiée, et $f : F[T_1, \ldots, T_r] \xrightarrow{\sim} \mathscr{A}$ un isomorphisme de F-algèbres. On a donc $f = ev_{\mathbf{a}}$, pour un unique $\mathbf{a} \in \mathscr{A}^r$ (cf. proposition 1.1). Mais alors, on a $\mathscr{A} = F[a_1, \ldots, a_r]$, puisque f est surjective. De plus, $a_1, \ldots, a_r$ sont algébriquement indépendants sur F par injectivité de f, d'où (1). □

On peut donc légitimement se poser la question de savoir si le résultat est encore vrai si l'on remplace $\mathfrak{S}_n$ par un sous-groupe G.

Autrement dit, soit G un sous-groupe de $\mathfrak{S}_n$, agissant sur $F[X_1, \ldots, X_n]$ par

$$\begin{aligned} G \times F[X_1, \ldots, X_n] &\longrightarrow F[X_1, \ldots, X_n] \\ (\sigma, P) &\longmapsto \sigma \cdot P = P(X_{\sigma(1)}, \ldots, X_{\sigma(n)}). \end{aligned}$$

Existe-t-il un isomorphisme de F-algèbres $F[X_1, \ldots, X_n]^G \simeq F[T_1, \ldots, T_r]$, pour un certain entier [2] $r \geqslant 0$?

La réponse est négative, comme le montre déjà le cas du groupe de Klein.

2.2. Exemple. Soient $s = (1\ 2)(3\ 4), t = (1\ 3)(2\ 4)$ et $G = \langle s, t \rangle \subset \mathfrak{S}_4$.

Soit F un corps de caractéristique différente de 2, et posons

$$U_1 = X_1 + X_2 + X_3 + X_4, \quad U_2 = X_1 - X_2 + X_3 - X_4,$$
$$U_3 = X_1 + X_2 - X_3 - X_4, \quad U_4 = X_1 - X_2 - X_3 + X_4.$$

Notons que

$$X_1 = \frac{U_1 + U_2 + U_3 + U_4}{4}, \quad X_2 = \frac{U_1 - U_2 + U_3 - U_4}{4},$$

$$X_3 = \frac{U_1 + U_2 - U_3 - U_4}{4}, \quad X_4 = \frac{U_1 - U_2 - U_3 + U_4}{4}.$$

Il est alors clair que $F[X_1, X_2, X_3, X_4] = F[U_1, U_2, U_3, U_4]$, et que $U_1, \ldots, U_4$ sont algébriquement indépendants sur F.

2. On peut en fait démontrer que, si c'est le cas, on a nécessairement $r = n$, mais encore une fois, nous n'aurons pas besoin de ce résultat.

On vérifie alors $U_2U_3U_4$, U_2^2, U_3^2, $U_4^2 \in \mathscr{A} = F[X_1, \ldots, X_4]^G$. Montrons que ces quatre éléments sont des éléments irréductibles de $\mathscr{A}$, deux à deux non associés. Comme $\mathscr{A}$ est un sous-anneau de $F[X_1, \ldots, X_4]$, $\mathscr{A}^\times$ est un sous-groupe de $F[X_1, \ldots, X_4]^\times = F^\times$. Comme les éléments de $F^\times$ sont clairement fixes sous l'action de G, on a $\mathscr{A}^\times = F^\times$.

Il est alors immédiat que $U_2U_3U_4, U_2^2, U_3^2, U_4^2$ sont deux à deux non associés. Pour vérifier qu'ils sont irréductibles, notons tout d'abord qu'ils sont non nuls et non inversibles. Montrons maintenant que $U_2U_3U_4$ est irréductible. Si $f \in \mathscr{A}$ est un diviseur de $U_2U_3U_4$ dans $\mathscr{A}$, c'est aussi un diviseur dans l'algèbre de polynômes $F[X_1, \ldots, X_4] = F[U_1, \ldots, U_4]$.

Comme $U_1, \ldots, U_4$ sont algébriquement indépendants, les diviseurs en question sont, à une constante multiplicative non nulle près :

$$1,\ U_2,\ U_3,\ U_4,\ U_2U_3,\ U_2U_4,\ U_3U_4,\ U_2U_3U_4.$$

Dans cette liste, seuls 1 et $U_2U_3U_4$ sont fixes sous l'action de G. Cela entraîne aisément l'irréductibilité de $U_2U_3U_4$.

On procède de même pour U_2^2, U_3^2, U_4^2. Mais alors, l'égalité

$$(U_2U_3U_4)^2 = U_2^2U_3^2U_4^2$$

montre que $(U_2U_3U_4)^2$ possède deux décompositions en irréductibles distinctes, et donc que l'anneau $\mathscr{A}$ n'est pas factoriel. En particulier, $\mathscr{A}$ ne peut être isomorphe à une algèbre de polynômes sur F.

On peut donc essayer de chercher une question un peu plus raisonnable. Au lieu de se préoccuper de polynômes, on peut s'intéresser aux fractions rationnelles.

Avant de continuer, on a besoin d'une définition.

2.3. Définition. Soit $\mathscr{A}$ une F-algèbre. Une action d'un groupe G sur $\mathscr{A}$ est une action *par automorphismes de F-algèbre* si pour tout $g \in G$, pour tous $a, a_1, a_2 \in \mathscr{A}$, et pour tout $\lambda \in F$, on a

(1) $g{\cdot}1_{\mathscr{A}} = 1_{\mathscr{A}}$

(2) $g{\cdot}(a_1 + a_2) = g{\cdot}a_1 + g{\cdot}a_2$

(3) $g{\cdot}(\lambda a) = \lambda(g{\cdot}a)$

(4) $g{\cdot}(a_1a_2) = (g{\cdot}a_1)(g{\cdot}a_2)$.

Autrement dit, une action d'un groupe G sur $\mathscr{A}$ est une action *par automorphismes de F-algèbre* si pour tout $g \in G$, la permutation

$$\begin{aligned} \mathscr{A} &\longrightarrow \mathscr{A} \\ a &\longmapsto g{\cdot}a \end{aligned}$$

est un automorphisme de F-algèbre, ou encore si le morphisme de groupes associé à l'action est en fait un morphisme $G \longrightarrow \mathrm{Aut}_{F-alg}(\mathscr{A})$.

2.4. Exemple. L'action d'un sous-groupe G de $\mathfrak{S}_n$ sur $F[X_1, \ldots, X_n]$ décrite précédemment est une action par automorphisme de F-algèbre.

On a alors le lemme suivant.

2.5. Lemme. *Soit G un groupe agissant sur $F[X_1, \ldots, X_n]$ par automorphismes de F-algèbre*[3].

Alors, l'application

$$\begin{aligned} G \times F(X_1, \ldots, X_n) &\longrightarrow F(X_1, \ldots, X_n) \\ \left(g, \frac{P}{Q}\right) &\longmapsto g \cdot \frac{P}{Q} = \frac{g \cdot P}{g \cdot Q} \end{aligned}$$

est bien définie, et est une action de G sur $F(X_1, \ldots, X_n)$ par automorphismes de F-algèbre.

Si de plus G est fini, on a

$$F(X_1, \ldots, X_n)^G = \left\{ \frac{P}{Q} \mid P \in F[X_1, \ldots, X_n]^G, Q \in F[X_1, \ldots, X_n]^G \backslash \{0\} \right\}.$$

Démonstration. Si $\dfrac{P_1}{Q_1} = \dfrac{P_2}{Q_2}$, on a $P_1 Q_2 = Q_1 P_2$. Puisque G agit par automorphismes d'algèbre, en appliquant l'action de g à cette égalité, on obtient $(g \cdot P_1)(g \cdot Q_2) = (g \cdot Q_1)(g \cdot P_2)$, et ainsi

$$\frac{g \cdot P_1}{g \cdot Q_1} = \frac{g \cdot P_2}{g \cdot Q_2}.$$

Cela montre que l'application de l'énoncé est bien définie. Le fait que cela définit une action par automorphismes de F-algèbre découle alors des définitions, et de vérifications un peu longues mais faciles.

Supposons maintenant que G soit fini, et montrons l'égalité annoncée. L'inclusion « $\supset$ » étant claire au vu des définitions, montrons l'autre inclusion. Soit $f = \dfrac{P}{Q} \in F(X_1, \ldots, X_n)^G$.

Posons $S = \prod_{g \in G} (g \cdot Q)$. On a alors $f = \dfrac{R}{S}$, pour un certain $R \in F[X_1, \ldots, X_n]$.

Remarquons que, par construction, on a $S \in F[X_1, \ldots, X_n]^G$. Montrons

3. Attention ! L'action considérée ici n'est pas nécessairement celle étudiée jusqu'à présent.

que $R \in F[X_1, \ldots, X_n]^G$. Pour cela, soit $g \in G$. On a alors

$$g \cdot f = \frac{g \cdot R}{g \cdot S} = \frac{g \cdot R}{S} = f = \frac{R}{S},$$

d'où $g \cdot R = R$. Cela étant vrai pour tout $g \in G$, on a bien $R \in F[X_1, \ldots, X_n]^G$, d'où le résultat. □

La seconde partie de ce lemme montre en particulier l'égalité

$$F(X_1, \ldots, X_n)^{\mathfrak{S}_n} = F(\sigma_1, \ldots, \sigma_n).$$

De plus, l'évaluation $T_i \rightsquigarrow \sigma_i$ induit un isomorphisme de F-algèbres

$$F(T_1, \ldots, T_n) \simeq F(X_1, \ldots, X_n)^{\mathfrak{S}_n}.$$

Comme précédemment, on a l'équivalence suivante, dont la démonstration est semblable en tous points à celle du lemme 2.1, mais cette fois en s'appuyant sur les résultats de la proposition 1.3.

2.6. Lemme. *Soit L/F une extension de corps et soit $r \geqslant 0$ un entier. Les propriétés suivantes sont équivalentes :*

(1) *il existe des éléments $a_1, \ldots, a_r \in L$ algébriquement indépendants sur F tels que $L = F(a_1, \ldots, a_r)$.*
(2) *il existe un isomorphisme de F-algèbres $F(T_1, \ldots, T_r) \simeq L$.*

Une question peut-être plus raisonnable serait donc la suivante.

Soit G un sous-groupe de $\mathfrak{S}_n$, agissant sur $F[X_1, \ldots, X_n]$ par

$$\begin{aligned} G \times F[X_1, \ldots, X_n] &\longrightarrow F[X_1, \ldots, X_n] \\ (\sigma, P) &\longmapsto \sigma \cdot P = P(X_{\sigma(1)}, \ldots, X_{\sigma(n)}). \end{aligned}$$

Considérons alors l'action de G induite sur $F(X_1, \ldots, X_n)$.

Problème de Noether. Existe-t-il un isomorphisme de F-algèbres

$$F(X_1, \ldots, X_n)^G \simeq F(T_1, \ldots, T_r),$$

pour un certain entier[4] $r \geqslant 0$?

Remarquons que si $F[X_1, \ldots, X_n]^G \simeq F[T_1, \ldots, T_r]$, on a aussi

$$F(X_1, \ldots, X_n)^G \simeq F(T_1, \ldots, T_n).$$

Il suffit en effet de chiosir des éléments $a_1, \ldots, a_r$ algébriquement indépendants sur F tels que $F[X_1, \ldots, X_n]^G = F[a_1, \ldots, a_r]$ et de considérer l'évaluation $T_i \rightsquigarrow a_i$.

4. Encore une fois, on peut en fait démontrer que, si c'est le cas, on a nécessairement $r = n$.

En revanche, la réciproque est fausse. Par exemple, on démontrera plus loin que, dans la situation de l'exemple 2.2, on a

$$F(X_1,\dots,X_4)^G = F(U_1,U_2^2,U_3^2,U_2U_3U_4),$$

et que U_1, U_2^2, U_3^2, $U_2U_3U_4$ sont algébriquement indépendants sur F.

Il est beaucoup plus difficile de donner des exemples de groupe G pour lequel le problème de Noether possède une réponse négative sans utiliser des outils sophistiqués. Mentionnons sans démonstration que lorsque $F=\mathbb{Q}$, le groupe $G=\mathbb{Z}/8\mathbb{Z}$ fournit un contre-exemple. En revanche, nous verrons que si $F=\mathbb{C}$, le problème de Noether admet une réponse positive pour tout groupe abélien fini. La réponse au problème de Noether dépend donc très fortement du corps F, ce qui rend cette question très délicate.

Nous allons maintenant donner quelques généralisations naturelles de cette question, qui sont considérées dans la littérature.

Une première généralisation facile est la suivante. On considère un groupe fini abstrait G, et l'on considère l'algèbre de polynômes $F[X_g, g\in G]$. Le groupe $\mathfrak{S}(G)$ agit sur $F[X_g,g\in G]$ par

$$\sigma\cdot P(X_g,g\in G) = P(X_{\sigma(g)},g\in G).$$

On peut alors identifier G à un sous-groupe de $\mathfrak{S}(G)$ grâce au plongement de Cayley, et faire agir G via ce sous-groupe. Autrement dit, on fait agir G sur $F[X_g\in G]$ par

$$g\cdot P(X_h,h\in G) = P(X_{gh},h\in H).$$

Cette action, ainsi que celle étudiée précédemment, peuvent être vues dans un contexte plus général. Avant d'expliquer lequel, on a besoin d'un lemme calculatoire.

Si $M=(a_{ij})\in \mathrm{M}_n(F)$, et si $P\in F[X_1,\dots,X_n]$, on pose

$$M*P(X_1,\dots,X_n) = P\Big(\sum_k a_{k1}X_k,\dots,\sum_k a_{kn}X_k\Big).$$

On a alors le résultat suivant.

2.7. Lemme. *Pour tous $M,N\in \mathrm{M}_n(F)$, et pour tout $P\in F[X_1,\dots,X_n]$, on a*

(1) $\mathrm{I}_n * P = P$

(2) $M*(N*P) = MN*P$.

Démonstration. Le premier point est clair. Montrons le second point.

Si $M=(a_{ij})$ and $N=(b_{ij})$, on a $MN=\big(\sum_k a_{ik}b_{kj}\big)$.

Posons $Q = N*P = P\big(\sum_k b_{k1}X_k,\dots,\sum_k b_{kn}X_k\big)$.

On a alors

$$\begin{aligned} M*(N*P) &= Q\Big(\sum_\ell a_{\ell 1}X_\ell,\dots,\sum_\ell a_{\ell n}X_\ell\Big) \\ &= P\Big(\sum_k b_{k1}\Big(\sum_\ell a_{\ell k}X_\ell\Big),\dots,\sum_k b_{kn}\Big(\sum_\ell a_{\ell k}X_\ell\Big)\Big). \end{aligned}$$

On a donc

$$M*(N*P)=P\Big(\sum_\ell\Big(\sum_k a_{\ell k}b_{k1}\Big)X_\ell\Big),\dots,\sum_\ell\Big(\sum_k a_{\ell k}b_{kn}\Big)X_\ell\Big)=MN*P.$$

□

2.8. Remarque. Par définition, on a l'égalité

$$\begin{pmatrix} M*X_1 \\ \vdots \\ M*X_n \end{pmatrix} = M^t \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}.$$

On peut alors définir toute une famille d'actions par automorphismes.

2.9. Lemme. *Soit G un groupe, et soit $\rho : G \longrightarrow \mathrm{GL}_n(F)$ un morphisme de groupes. Alors, l'application*

$$\begin{aligned} G\times F[X_1,\dots,X_n] &\longrightarrow F[X_1,\dots,X_n] \\ (g,P) &\longmapsto g\cdot P=\rho(g)^{-t}*P \end{aligned}$$

définit une action de G sur $F[X_1,\dots,X_n]$ par automorphismes de F-algèbres.

Démonstration. Ce sont de simples calculs. □

2.10. Exemples

(1) Pour tout $\sigma\in\mathfrak{S}_n$, on pose $M_\sigma=(\delta_{i\sigma(j)})$.

Pour tout sous-groupe G de $\mathfrak{S}_n$, l'application

$$\begin{aligned} \rho : G &\longrightarrow \mathrm{GL}_n(F) \\ \sigma &\longmapsto M_\sigma \end{aligned}$$

est alors un morphisme de groupes. En particulier, $M_\sigma^{-t}=M_{\sigma^{-1}}^t=(\delta_{j\sigma^{-1}(i)})$. On a alors $\sum_k \delta_{j\sigma^{-1}(k)}X_k = X_{\sigma(j)}$, et pour tout $P\in F[X_1,\dots,X_n]$, on a donc $\sigma\cdot P = P(X_{\sigma(1)},\dots,X_{\sigma(n)})$.

(2) Soit G un groupe d'ordre n. On indexe les éléments de $\mathrm{M}_n(F)$ par les éléments de G. Pour tout $g\in G$, on pose

$$\rho'(g)=(\delta_{h,gk})_{h,k\in G}.$$

L'application $\rho' : G\longrightarrow \mathrm{GL}_n(F)$ est alors un morphisme de groupes. C'est en fait le morphisme obtenu en composant le plongement de Cayley et le morphisme ρ du point précédent.

Il est alors facile de constater que pour tout $g \in G$ et tout $P \in F[X_h, h \in G]$, on a $g{\cdot}P = P(X_{gh}, h \in G)$.

On peut maintenant utiliser la famille d'actions par automorphisme définie dans le lemme 2.9 pour généraliser les questions précédentes.

Soit $\rho : G \longrightarrow \mathrm{GL}_n(F)$ un morphisme de groupes. On considère les actions de G induites sur $F[X_1, \ldots, X_n]$ et $F(X_1, \ldots, X_n)$.

(i) La F-algèbre $F[X_1, \ldots, X_n]^G$ est-elle isomorphe à une algèbre de polynômes sur F ?

(ii) **Problème de Noether généralisé.** La F-algèbre $F(X_1, \ldots, X_n)^G$ est-elle isomorphe à une algèbre de fractions rationnelles sur F ?

Comme on l'a déjà constaté, la réponse à la question (i) est déjà négative pour le groupe de Klein (et positive pour (ii)), mais avec cette version généralisée, on peut donner un contre-exemple avec $G = \mathbb{Z}/2\mathbb{Z}$.

2.11. Exemple. Soit F un corps de caractéristique différente de 2.

Soit $\rho : \mathbb{Z}/2\mathbb{Z} \longrightarrow \mathrm{GL}_2(F)$ l'unique morphisme envoyant $\overline{1}$ sur $-\mathrm{I}_2$. L'action correspondante sur $F[X_1, X_2]$ est donnée par

$$\overline{m}{\cdot}P \;=\; P\big((-1)^m X_1, (-1)^m X_2\big).$$

Il est facile de voir [5] que $F[X_1, X_2]^{\mathbb{Z}/2\mathbb{Z}} = \mathrm{Vect}_F(X_1^m X_2^n, m \equiv n\ [2])$. Remarquons que si $m \equiv n\ [2]$, on a $m = n + 2k$ ou $n = m + 2k, k \geqslant 0$, selon que $m \geqslant n$ ou $n \geqslant m$. Dans le premier cas, on a $X_1^m X_2^n = (X_1X_2)^n (X_1^2)^k$ et $X_1^m X_2^n = (X_1X_2)^m (X_2^2)^k$ dans le second.

On en déduit rapidement l'égalité $F[X_1, X_2]^{\mathbb{Z}/2\mathbb{Z}} = F[X_1^2, X_1X_2, X_2^2]$.

Comme dans l'exemple 2.2, on montre que X_1^2, X_1X_2 et X_2^2 sont irréductibles, non associés deux à deux et que $F[X_1^2, X_1X_2, X_2^2]$ ne peut être isomorphe à une algèbre de polynômes sur F.

En revanche, la réponse au problème de Noether dans ce cas est encore positive.

En effet, le lemme 2.5 montre que $F(X_1, X_2)^{\mathbb{Z}/2\mathbb{Z}} = F(X_1^2, X_1X_2, X_2^2)$.

Comme $X_1^2 = \dfrac{(X_1X_2)^2}{X_2^2}$, on a $F(X_1, X_2)^{\mathbb{Z}/2\mathbb{Z}} = F(X_1X_2, X_2^2)$.

Puisque X_1X_2 et X_2^2 sont algébriquement indépendants sur F, on a bien un isomorphisme $F(X_1, X_2)^{\mathbb{Z}/2\mathbb{Z}} \simeq F(T_1, T_2)$.

On peut se demander pourquoi définir l'action de G sur $F[X_1, \ldots, X_n]$ de manière si compliquée, au lieu de simplement poser $g{\cdot}P = \rho(g) * P$.

5. Cela utilise le fait que $-1 \neq 1$ dans F, d'où l'hypothèse de caractéristique sur F.

L'action proposée est en fait la plus naturelle. Pour l'expliquer, nous allons réinterpréter les actions de G sur $F[X_1, \ldots, X_n]$ et $F(X_1, \ldots, X_n)$ dans le contexte de la théorie des représentations.

Introduisons une notation utile.

Notation. Soit $\mathscr{A}$ une F-algèbre **commutative**, et soit S une partie (éventuellement vide) de $\mathscr{A}$. On note

$$F[S] = \{P(\mathbf{s}) \mid n \geqslant 0,\ \mathbf{s} \in S^n,\ P \in F[X_1, \ldots, X_n]\}.$$

C'est en fait la sous-algèbre de $\mathscr{A}$ engendrée par S.

On peut maintenant définir la notion de fonction polynomiale sur un espace vectoriel.

2.12. Définition. Soit V un F-espace vectoriel de dimension finie $n \geqslant 1$.

Une *fonction monomiale sur* V est une fonction $V \longrightarrow F$ de la forme

$$c\varphi_1 \cdots \varphi_r,$$

où $r \geqslant 0$, $c \in F^\times$ et $\varphi_1, \ldots, \varphi_r \in V^*$.

Une *fonction polynomiale sur* V est une somme finie (éventuellement vide) de fonctions monomiales sur V.

L'ensemble des fonctions polynomiales sur V est une sous-algèbre commutative de l'algèbre $\mathscr{F}(V, F)$ des fonctions de V dans F.

Ce n'est rien d'autre que la sous-algèbre $F[V^*]$ de $\mathscr{F}(V, F)$ engendrée par les éléments de V^*.

Explicitons un peu plus la définition. Soit $(e_1, \ldots, e_n)$ une F-base de V, et soit $(e_1^*, \ldots, e_n^*)$ la base duale de V^*. Toute forme linéaire sur V étant combinaison linéaire de $e_1^*, \ldots, e_n^*$, on constate facilement qu'une fonction polynomiale sur V est une combinaison linéaire de fonctions monomiales de la forme $(e_1^*)^{m_1} \cdots (e_n^*)^{m_n}$, $m_i \geqslant 0$.

Autrement dit, une fonction polynomiale $f : V \longrightarrow F$ s'écrit sous la forme

$$f = \sum_{\mathbf{m}} \lambda_{\mathbf{m}} (e_1^*)^{m_1} \cdots (e_n^*)^{m_n}, \quad \lambda_{\mathbf{m}} \in F.$$

On a alors $f\Big(\sum_{i=1}^{n} x_i e_i\Big) = \sum_{\mathbf{m}} \lambda_{\mathbf{m}} x_1^{m_1} \cdots x_n^{m_n}$ pour tout $x_i \in F$, ce qui justifie le nom de fonction polynomiale.

Notons que l'on a un morphisme d'évaluation

$$\begin{aligned} F[X_1, \ldots, X_n] &\longrightarrow F[V^*] \\ P &\longmapsto P(e_1^*, \ldots, e_n^*) \end{aligned}$$

qui est surjectif, d'après ce qui précède.

En général, ce morphisme n'est pas injectif, comme on le voit déjà dans le cas $F = \mathbb{F}_2$, puisque le polynôme $X^2 - X$ s'évalue en la fonction nulle.

En revanche, si F est infini, tout se passe bien.

2.13. Lemme. *Soit V un F-espace vectoriel de dimension $n \geqslant 1$, où F est un corps infini, et soit $(e_1, \ldots, e_n)$ une base de V. Alors, le morphisme d'évaluation*

$$\begin{aligned} F[X_1, \ldots, X_n] &\longrightarrow F[V^*] \\ P &\longmapsto P(e_1^*, \ldots, e_n^*) \end{aligned}$$

est un isomorphisme de F-algèbres.

Démonstration. Il reste à démontrer l'injectivité du morphisme d'évaluation. Cela revient à démontrer le fait suivant : si F est un corps infini, alors pour tout $P \in F[X_1, \ldots, X_n]$, on a

$$P(x_1, \ldots, x_n) = 0, \quad \text{pour tous } x_1, \ldots, x_n \in F \implies P = 0.$$

C'est un résultat bien connu pour $n = 1$ (un polynôme en une indéterminée ayant une infinité de racines est nécessairement nul), et le cas général découle d'une récurrence finie. □

Nous pouvons maintenant expliquer d'où vient l'action définie dans le lemme 2.9. Rappelons que, si G agisse linéairement sur V, alors G agit linéairement sur $\mathscr{F}(V, F)$ par

$$(g \cdot f)(v) = f(g^{-1} \cdot v), \quad \text{pour tout } v \in V.$$

De plus, si $\varphi \in V^*$, on a aussi $g \cdot \varphi \in V^*$ pour tout $g \in G$. Il s'ensuit que cette action de G sur $\mathscr{F}(V, F)$ se restreint en une action linéaire sur V^*, et en une action par automorphismes d'algèbre sur $F[V^*]$.

Soit $\mathscr{B} = (e_1, \ldots, e_n)$ une base de V, et soit $\mathscr{B}^* = (e_1^*, \ldots, e_n^*)$ la base duale de V^*.

Si l'on écrit les éléments de V dans la base $\mathscr{B}$ l'action de G sur V correspond à un morphisme $\rho : G \longrightarrow \mathrm{GL}_n(F)$. Le morphisme $\rho^* : G \longrightarrow \mathrm{GL}_n(F)$ correspondant à l'action induite sur V^* est alors défini par

$$\rho^*(g) = \rho(g)^{-t}, \quad \text{pour tout } g \in G,$$

d'après l'exemple XXXIV-4.4.

Autrement dit, si $\rho(g)^{-t} = (a'_{ij})$, alors $g \cdot e_j^* = \sum_k a'_{kj} e_k^*$.

Si $f = P(e_1^*, \ldots, e_n^*) \in F[V^*]$, on a donc

$$g \cdot f = P(g \cdot e_1^*, \ldots, g \cdot e_n^*) = P\Big(\sum_k a'_{k1} e_k^*, \ldots, \sum_k a'_{kn} e_k^*\Big),$$

la deuxième égalité provenant du fait que G agit par automorphismes d'algèbre. Lorsque F est infini, en transportant l'action de G sur $F[V^*]$ en

une action de G sur $F[X_1, \ldots, X_n]$ via l'isomorphisme du lemme 2.13, on obtient alors l'action définie dans le lemme 2.9.

3. Le théorème de Fischer : aspect théorique

Nous allons maintenant étudier le problème de Noether pour les groupes abéliens finis lorsque le corps de base contient suffisamment de racines de l'unité. Plus précisément, nous allons montrer le théorème suivant.

3.1. Théorème (Fischer). *Soit G un groupe abélien fini d'exposant e, et soit F un corps tel que* $\mathrm{car}(F) \nmid e$ *et* $\mu_e(\overline{F}) \subset F$.

Alors, pour tout morphisme $\rho : G \longrightarrow \mathrm{GL}_n(F)$, on a un isomorphisme d'algèbres

$$F(X_1, \ldots, X_n)^G \simeq F(T_1, \ldots, T_n).$$

Autrement dit, il existe $f_1, \ldots, f_n \in F(X_1, \ldots, X_n)^G$ algébriquement indépendantes sur F telles que

$$F(X_1, \ldots, X_n)^G = F(f_1, \ldots, f_n).$$

3.2. Remarque. Ce théorème n'est plus vrai si l'on remplace $F(X_1, \ldots, X_n)$ par $F[X_1, \ldots, X_n]$, comme le montre déjà l'exemple 2.11.

Nous allons maintenant nous employer à démontrer le théorème de Fischer. Dans cette partie, nous fournirons une démonstration théorique de ce résultat, mais nous verrons dans la partie suivante une méthode explicite pour tout calculer.

Dans la suite, on fixe un groupe abélien fini G d'exposant e, et un morphisme $\rho : G \longrightarrow \mathrm{GL}_n(F)$, où F vérifie $\mathrm{car}(F) \nmid e$ et $\mu_e(\overline{F}) \subset F$.

On commence par faire un changement de variables qui sera plus adapté pour nos calculs.

3.3. Lemme

(1) *Il existe une matrice inversible $R \in \mathrm{GL}_n(F)$ et des morphismes de groupes $\omega_i : G \longrightarrow \mu_e(\overline{F})$ (pas nécessairement distincts) tels que*

$$R^{-1}\rho(g)R = \begin{pmatrix} \omega_1(g) & & \\ & \ddots & \\ & & \omega_n(g) \end{pmatrix}, \quad \textit{pour tout } g \in G.$$

(2) *Soit $R \in \mathrm{GL}_n(F)$ une matrice vérifiant la propriété précédente.*

Si $R^{-1} = (a'_{ij})$, *on pose*

$$U_j = \sum_k a'_{jk} X_k, \quad \textit{pour tout } j \in [\![1, n]\!].$$

Alors, $F[X_1, \ldots, X_n] = F[U_1, \ldots, U_n]$, *et* $U_1, \ldots, U_n$ *sont algébriquement indépendants sur* F. *De plus,*

$$g{\cdot}U_i = \omega_i(g)^{-1} U_i, \quad \textit{pour tout } g \in G, \textit{ et tout } i \in [\![1, n]\!].$$

Démonstration. Soit $g \in G$. Puisque $g^e = 1_G$, on a $\rho(g)^e = \rho(g^e) = \mathrm{I}_n$. Les hypothèses sur F entraînent que $X^e - 1$ est scindé sur F à racines simples. Ainsi, $\rho(g)$ est diagonalisable, et ses valeurs propres sont des racines e-ièmes de l'unité. Puisque G est abélien, les matrices $\rho(g), g \in G$, sont donc codiagonalisables [6].

Il existe donc $R \in \mathrm{GL}_n(F)$ tel que

$$R^{-1}\rho(g)R = \begin{pmatrix} \omega_1(g) & & \\ & \ddots & \\ & & \omega_n(g) \end{pmatrix}, \quad \text{pour tout } g \in G,$$

où $\omega_1(g), \ldots, \omega_n(g) \in \mu_e(\overline{F})$.

Puisque ρ est un morphisme de groupes, les applications $\omega_i : G \longrightarrow \mu_e(\overline{F})$ sont des morphismes de groupes, comme on le voit grâce à un simple calcul. Cela démontre le point (1).

Démontrons (2). Par définition de $U_1, \ldots, U_n$, on a $\begin{pmatrix} U_1 \\ \vdots \\ U_n \end{pmatrix} = R^{-1} \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix}$.

Puisque $g{\cdot}U_i = \sum_j a'_{ij} g{\cdot}X_j$, on a donc $\begin{pmatrix} g{\cdot}U_1 \\ \vdots \\ g{\cdot}U_n \end{pmatrix} = R^{-1} \begin{pmatrix} g{\cdot}X_1 \\ \vdots \\ g{\cdot}X_n \end{pmatrix}$.

D'après la remarque 2.8 et la définition de l'action de G sur $F[X_1, \ldots, X_n]$, on a

$$\begin{pmatrix} g{\cdot}X_1 \\ \vdots \\ g{\cdot}X_n \end{pmatrix} = (\rho(g)^{-t})^t \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = \rho(g)^{-1} \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix},$$

si bien que

$$\begin{pmatrix} g{\cdot}U_1 \\ \vdots \\ g{\cdot}U_n \end{pmatrix} = R^{-1}\rho(g)^{-1} \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} = R^{-1}\rho(g)^{-1} R \begin{pmatrix} U_1 \\ \vdots \\ U_n \end{pmatrix},$$

6. Cela reflète le fait que les représentations irréductibles d'un groupe abélien sont toutes de degré 1.

soit encore

$$\begin{pmatrix} g{\cdot}U_1 \\ \vdots \\ g{\cdot}U_n \end{pmatrix} = \begin{pmatrix} \omega_1(g) & & \\ & \ddots & \\ & & \omega_n(g) \end{pmatrix}^{-1} \begin{pmatrix} U_1 \\ \vdots \\ U_n \end{pmatrix}.$$

Ainsi, on obtient

$$g{\cdot}U_i = \omega_i(g)^{-1}U_i, \quad \text{pour tout } g \in G, \text{ et tout } i \in [\![1, n]\!].$$

Notons maintenant que, par définition des U_i, on a

$$P(U_1, \ldots, U_n) = R^{-t} * P(X_1, \ldots, X_n), \quad \text{pour tout } P \in F[X_1, \ldots, X_n].$$

Comme $P = R^t * (R^{-t} * P) = R^t * P(U_1, \ldots, U_n)$, on en déduit l'égalité

$$F[U_1, \ldots, U_n] \;=\; F[X_1, \ldots, X_n].$$

Enfin, si $Q \in F[T_1, \ldots, T_n]$ vérifie $Q(U_1, \ldots, U_n) = 0$, on a $R^{-t} * Q = 0$, ce qui entraîne $Q = R^t * (R^{-t} * Q) = 0$. Ainsi, $U_1, \ldots, U_n$ sont algébriquement indépendants sur F. □

Nous pouvons maintenant décrire une première famille de générateurs de l'algèbre des polynômes et fractions rationnelles G-invariantes.

Avant d'énoncer le résultat, on introduit une dernière notation, qui étend une notation précédente.

Notation. Si $\boldsymbol{\gamma} = \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix} \in \mathbb{Z}^n$, on pose $T^{\boldsymbol{\gamma}} = T_1^{\gamma_1} \cdots T_n^{\gamma_n} \in F(T_1, \ldots, T_n)$.

On a alors le lemme suivant.

3.4. Lemme. *Soient*

$$\mathscr{M} = \Big\{ \mathbf{m} \in \mathbb{N}^n \;\Big|\; \omega_1(g)^{m_1} \cdots \omega_n(g)^{m_n} = 1, \;\; \textit{pour tout } g \in G \Big\}$$

et

$$\Gamma = \Big\{ \boldsymbol{\gamma} \in \mathbb{Z}^n \;\Big|\; \omega_1(g)^{\gamma_1} \cdots \omega_n(g)^{\gamma_n} = 1, \;\; \textit{pour tout } g \in G \Big\}.$$

Alors, on a les égalités suivantes :

(1) $F[X_1, \ldots, X_n]^G = \mathrm{Vect}_F(U^{\mathbf{m}}, \mathbf{m} \in \mathscr{M}) = F[U^{\mathbf{m}}, \mathbf{m} \in \mathscr{M}]$.

(2) $F(X_1, \ldots, X_n)^G = F(U^{\mathbf{m}}, \mathbf{m} \in \mathscr{M}) = F(U^{\boldsymbol{\gamma}}, \boldsymbol{\gamma} \in \Gamma)$.

Démonstration. Gardons les notations du lemme précédent. Par ce même lemme, on a $F[X_1, \ldots, X_n] = F[U_1, \ldots, U_n]$, où les U_i sont algébriquement indépendants sur F. Autrement dit, tout polynôme $P \in F[X_1, \ldots, X_n]$

s'écrit de manière unique sous la forme

$$P = \sum_{\mathbf{m}} \lambda_{\mathbf{m}} U^{\mathbf{m}}, \quad \lambda_{\mathbf{m}} \in F.$$

On a alors

$$g \cdot P = \sum_{\mathbf{m}} \lambda_{\mathbf{m}} \omega_1(g)^{-m_1} \cdots \omega_n(g)^{-m_n} U^{\mathbf{m}}, \quad \text{pour tout } g \in G.$$

L'égalité précédente montre que l'on a $P \in F[X_1, \ldots, X_n]^G$ si, et seulement si, $\lambda_{\mathbf{m}} \omega_1(g)^{-m_1} \cdots \omega_n(g)^{-m_n} = \lambda_{\mathbf{m}}$, pour tout $g \in G$ et tout $\mathbf{m} \in \mathbb{N}^n$.

On en déduit alors facilement l'égalité

$$F[X_1, \ldots, X_n]^G = \mathrm{Vect}_F(U^{\mathbf{m}},\ \mathbf{m} \in \mathscr{M}),$$

puis

$$F[X_1, \ldots, X_n]^G = F[U^{\mathbf{m}},\ \mathbf{m} \in \mathscr{M}].$$

En effet, l'inclusion $\mathrm{Vect}_F(U^{\mathbf{m}}, \mathbf{m} \in \mathscr{M}) \subset F[U^{\mathbf{m}}, \mathbf{m} \in \mathscr{M}]$ est évidente, et l'autre inclusion provient du fait qu'un produit fini de monômes de la forme $U^{\mathbf{m}}$, $\mathbf{m} \in \mathscr{M}$, est encore un monôme de cette forme, car $\mathscr{M}$ est stable par addition.

Le lemme 2.5 implique alors l'égalité

$$F(X_1, \ldots, X_n)^G = F(U^{\mathbf{m}},\ \mathbf{m} \in \mathscr{M}).$$

Montrons (2). Encore une fois, l'inclusion

$$F(X_1, \ldots, X_n)^G = F(U^{\mathbf{m}},\ \mathbf{m} \in \mathscr{M}) \subset F(U^{\gamma},\ \gamma \in \Gamma)$$

est évidente.

Pour établir l'autre inclusion, il suffit de voir que $U^{\gamma} \in F(U^{\mathbf{m}}, \mathbf{m} \in \mathscr{M})$ pour tout $\gamma \in \Gamma$.

Or, on a $e\mathbb{Z}^n \subset \Gamma$ par définition de l'exposant. Si $\gamma \in \Gamma$, choisissons $k \in \mathbb{N}$ tel que $m_i' = ke + \gamma_i \geqslant 0$, pour tout $i \in [\![1, n]\!]$.

Alors, $\begin{pmatrix} ke \\ \vdots \\ ke \end{pmatrix}$ et $\mathbf{m}'$ sont des éléments de $\mathscr{M}$, et l'on a

$$U^{\gamma} = \frac{U_1^{m_1'} \cdots U_n^{m_n'}}{U_1^{ke} \cdots U_n^{ke}} \in F(U^{\mathbf{m}},\ \mathbf{m} \in \mathscr{M}),$$

d'où la conclusion souhaitée. □

On peut se demander ce que l'on a gagné en remplaçant $\mathscr{M}$ par Γ dans le point (2), puisque, a priori, on a rajouté des générateurs inutiles. En fait, Γ étant un sous-groupe de $\mathbb{Z}^n$, il possède une $\mathbb{Z}$-base (cf. corollaire XII-4.11).

On peut alors démontrer le résultat suivant, qui permet de conclure.

3.5. Proposition. *Soit* $(\gamma^{(1)},\dots,\gamma^{(n)})$ *une* $\mathbb{Z}$*-base de* Γ*. Alors,*

(1) $F(X_1,\dots,X_n)^G = F\big(U^{\gamma^{(j)}},\ j\in[\![1,n]\!]\big)$;

(2) *les fractions rationnelles* $U^{\gamma^{(j)}}$, $j\in[\![1,n]\!]$, *sont algébriquement indépendantes sur* F.

Démonstration. On sait déjà que $F(X_1,\dots,X_n)^G = F(U^\gamma,\gamma\in\Gamma)$. Pour démontrer le point (1), il suffit donc de constater l'égalité

$$F(U^\gamma,\gamma\in\Gamma) = F\big(U^{\gamma^{(j)}},\ j\in[\![1,n]\!]\big).$$

L'inclusion « $\supset$ » est évidente. Démontrons l'inclusion manquante.

Soit $\gamma = \begin{pmatrix}\gamma_1\\ \vdots\\ \gamma_n\end{pmatrix} \in \Gamma$. Il existe donc $a_1,\dots,a_n\in\mathbb{Z}$ tels que

$$\gamma = a_1\gamma^{(1)}+\cdots+a_n\gamma^{(n)}.$$

On a alors

$$U^\gamma = \prod_{j=1}^n (U^{\gamma^{(j)}})^{a_j} \in F\big(U^{\gamma^{(j)}},\ j\in[\![1,n]\!]\big).$$

Cela suffit à conclure. Montrons le point (2).

Soit $P=\sum_{\mathbf{m}}\lambda_{\mathbf{m}}T^{\mathbf{m}}\in F[T_1,\dots,T_n]$ tel que

$$P\big(U^{\gamma^{(1)}},\dots,U^{\gamma^{(n)}}\big)=0.$$

On a donc

$$\sum_{\mathbf{m}}\lambda_{\mathbf{m}}U^{\sum_{j=1}^n m_j\gamma^{(j)}}=0.$$

Soit $\mathbf{u}=\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}$, et soit $k\geqslant 1$ un entier tel que $k\mathbf{u}+\sum_{j=1}^n m_j\gamma^{(j)}\in\mathbb{N}^n$, pour tout $\mathbf{m}\in\mathbb{N}^n$ tel que $\lambda_{\mathbf{m}}\neq 0$.

En multipliant l'égalité précédente par $(U_1\cdots U_n)^k$, on obtient

$$\sum_{\mathbf{m}}\lambda_{\mathbf{m}}U^{k\mathbf{u}+\sum_{j=1}^n m_j\gamma^{(j)}}=0\in F[U_1,\dots,U_n].$$

Montrons que les n-uplets $k\mathbf{u}+\sum_{j=1}^n m_j\gamma^{(j)}$, lorsque $\mathbf{m}$ parcourt $\mathbb{N}^n$, sont distincts deux à deux.

Supposons que

$$k\mathbf{u}+\sum_{j=1}^n m_j\gamma^{(j)} = k\mathbf{u}+\sum_{j=1}^n m'_j\gamma^{(j)}.$$

On a donc $\sum_{j=1}^n (m_j-m'_j)\gamma^{(j)}=0$, pour tout $i\in[\![1,n]\!]$. Comme $(\gamma^{(1)},\dots,\gamma^{(n)})$ est une $\mathbb{Z}$-base, on en déduit que $m_j=m'_j$, pour tout $j\in[\![1,n]\!]$, ce qu'il fallait montrer.

L'égalité

$$\sum_{\mathbf{m}} \lambda_{\mathbf{m}} U^{k\mathbf{u}+\sum_{j=1}^{n} m_j \boldsymbol{\gamma}^{(j)}} = 0 \in F[U_1, \ldots, U_n]$$

fait donc intervenir des monômes deux à deux distincts. On en déduit que tous les coefficients correspondants sont nuls, i.e. $\lambda_{\mathbf{m}} = 0$ pour tout $\mathbf{m} \in \mathbb{N}^n$. Ainsi, $P = 0$, et l'on a fini. □

4. Le théorème de Fischer : aspect pratique

Nous allons maintenant nous intéresser à l'aspect calculatoire du théorème de Fischer. Le but de cette partie est de donner une méthode pratique pour calculer des fractions rationnelles G-invariantes algébriques indépendantes engendrant $F(X_1, \ldots, X_n)^G$.

La démonstration du théorème de Fischer donnée dans la partie précédente nous fournit a priori un moyen de tout calculer à deux exceptions près :

(1) il faut réussir à exhiber une matrice R satisfaisant les conditions du lemme 3.3 ;

(2) il faut pouvoir trouver une $\mathbb{Z}$-base du groupe abélien Γ défini dans le lemme 3.4.

Intéressons-nous tout d'abord au calcul de R, c'est-à-dire au calcul d'une base de diagonalisation commune des matrices $\rho(g), g \in G$.

Le point (1) du lemme 3.3 nous fournit l'existence d'une base $(v_1, \ldots, v_n)$ de F^n telle que

$$\rho(g)(v_i) = \omega_i(g) v_i, \quad \text{pour tout } g \in G \text{ et tout } i \in [\![1, n]\!],$$

où $\omega_i : G \longrightarrow \mu_e(\overline{F})$ est un morphisme de groupes.

La première étape est de décrire ces morphismes. Pour éviter de rallonger l'exposition, on suppose que l'on a déjà à disposition une décomposition de la forme $G = \langle g_1 \rangle \odot \cdots \odot \langle g_r \rangle$ en produit direct de sous-groupes cycliques, d'ordres $d_1, \ldots, d_r$ respectivement (avec $2 \leqslant d_1 \mid \cdots \mid d_r$ si l'on souhaite avoir un r minimal).

Tout élément $g \in G$ s'écrit alors de manière unique sous la forme

$$g = g_1' \cdots g_r', \quad g_i' \in \langle g_i \rangle.$$

Chaque g_i' s'écrit $g_i' = g_i^{m_i}, m_i \in \mathbb{Z}$, où m_i est unique modulo d_i.

On a alors le lemme suivant.

4.1. Lemme. *Pour chaque $i \in [\![1,r]\!]$, fixons $z_i \in \mu_{d_i}(\overline{F})$. Alors, il existe un unique morphisme de groupes $\omega_{z_1,\dots,z_n} : G \longrightarrow \mu_e(\overline{F})$ tel que*

$$\omega_{z_1,\dots,z_n}(g_i) = z_i, \quad \textit{pour tout } i \in [\![1,r]\!].$$

Il est défini par

$$\omega_{z_1,\dots,z_n}(g_1^{m_1}\cdots g_r^{m_r}) = z_1^{m_1}\cdots z_r^{m_r}, \quad m_i \in \mathbb{Z}.$$

De plus, tout morphisme $\omega : G \longrightarrow \mu_e(\overline{F})$ est de la forme $\omega_{z_1,\dots,z_n}$, où les racines de l'unité $z_1,\dots,z_n$ sont uniques.

Démonstration. Notons que $d_i \mid e$, puisque $g_i^e = 1_G$, et donc que z_i est bien une racine e-ième de l'unité. On remarque ensuite que $\omega_{z_1,\dots,z_n}(g_1^{m_1}\cdots g_r^{m_r})$ ne dépend pas du choix des m_i, puisque $z_i^{d_i} = 1$. Il est alors facile de voir que $\omega_{z_1,\dots,z_n}$ est un morphisme de groupes.

Soit maintenant $\omega : G \longrightarrow \mu_e(\overline{F})$ un morphisme de groupes. Notons que

$$\omega(g_i)^{d_i} \;=\; \omega(g_i^{d_i}) \;=\; \omega(1_G) = 1.$$

On a donc nécessairement $\omega(g_i) \in \mu_{d_i}(\overline{F})$. Si l'on pose $z_i = \omega(g_i)$, on a donc

$$\omega(g_1^{m_1}\cdots g_r^{m_r}) = z_1^{m_1}\cdots z_r^{m_r}, \quad \text{pour tous } m_1,\dots,m_r \in \mathbb{Z},$$

et ainsi $\omega = \omega_{z_1,\dots,z_r}$.

Si maintenant $\omega = \omega_{z'_1,\dots,z'_r}$, alors en évaluant les images en chaque g_i, on obtient $z'_i = \omega(g_i)$, d'où la partie unicité. □

On en déduit un résultat qui nous sera fort utile un peu plus loin.

4.2. Corollaire. *Soit G un groupe abélien fini d'exposant e.*

Pour tous morphismes de groupes $\omega, \omega' : G \longrightarrow \mu_e(\overline{F})$, on a

$$\frac{1}{|G|}\sum_{g\in G}\omega(g)^{-1}\omega'(g) = \begin{cases} 1 & \textit{si } \omega = \omega' \\ 0 & \textit{sinon.} \end{cases}$$

Démonstration. Le cas $\omega = \omega'$ étant clair, on peut supposer que $\omega \neq \omega'$.

Gardons les notations du lemme précédent. D'après ce même lemme, on a

$$\omega = \omega_{z_1,\dots,z_r} \quad \text{et} \quad \omega' = \omega_{z'_1,\dots,z'_r},$$

avec $z_i, z'_i \in \mu_{d_i}(\overline{F})$. On a donc $z_j \neq z'_j$, pour un certain $j \in [\![1,r]\!]$.

Notons qu'un élément $g \in G$ s'écrit de manière unique

$$g = g_1^{k_1}\cdots g_r^{k_r}, \quad k_i \in [\![0, d_i - 1]\!].$$

On a alors

$$\frac{1}{|G|}\sum_{g\in G}\omega(g)^{-1}\omega'(g)=\frac{1}{|G|}\Big(\sum_{k_1=0}^{d_1-1}(z_1^{-1}z_1')^{k_1}\Big)\cdots\Big(\sum_{k_r=0}^{d_r-1}(z_r^{-1}z_r')^{k_r}\Big).$$

Or, si z est une racine d-ième de l'unité différente de 1, on a

$$\sum_{k=0}^{d-1}z^k=\frac{1-z^d}{1-z}=0.$$

Par hypothèse, le j-ème facteur de l'expression précédente est donc nul, puisque $z_j^{-1}z_j'\neq 1$. Cela achève la démonstration. □

Reprenons nos considérations. Si $\omega : G \longrightarrow \mu_e(\overline{F})$ est un morphisme de groupes, on pose

$$V_\omega=\bigcap_{g\in G}\operatorname{Ker}\big(\rho(g)-\omega(g)\mathrm{Id}\big)=\big\{v\in F^n\,|\,\rho(g)v=\omega(g)v,\ \text{pour tout } g\in G\big\}.$$

Le lemme 3.3 nous dit alors que l'on a une décomposition en somme directe

$$F^n=\bigoplus_{\omega}V_\omega,$$

où ω décrit l'ensemble des morphismes $G \longrightarrow \mu_e(\overline{F})$. Notons que V_ω peut très bien être le sous-espace nul. Pour éviter cela, on peut bien sûr limiter la somme aux seuls morphismes aux seuls morphismes ω_i qui apparaissent dans le lemme 3.3, mais cela suppose de factoriser les polynômes caractéristiques des matrices $\rho(g)$, $g \in G$ (ou au moins des $\rho(g_k)$).

Tout le jeu consiste maintenant à calculer une base de V_ω pour chaque morphisme ω à partir de la seule connaissance de ρ. La solution est donnée par la proposition suivante.

4.3. Proposition. *Pour tout morphisme $\omega : G \longrightarrow \mu_e(\overline{F})$, on pose*

$$P_\omega=\frac{1}{|G|}\sum_{g\in G}\omega(g)^{-1}\rho(g).$$

Alors, P_ω est la matrice dans la base canonique de la projection sur V_ω parallèlement à $\bigoplus_{\omega'\neq\omega}V_{\omega'}$.

Démonstration. Soit $\omega' : G \longrightarrow \mu_e(\overline{F})$ un morphisme, et soit $v \in V_{\omega'}$. On doit démontrer que $P_\omega v = v$ si $\omega = \omega'$ et que $P_\omega v = 0$ si $\omega \neq \omega'$.

Par définition de $V_{\omega'}$, on a

$$P_\omega v = \frac{1}{|G|} \sum_{g \in G} \omega(g)^{-1} \omega'(g) v.$$

On utilise alors le lemme précédent pour conclure. □

Autrement dit, pour calculer une base de diagonalisation simultanée, il suffit de procéder aux étapes suivantes :

(1) On calcule tous les morphismes $\omega : G \longrightarrow \mu_e(\overline{F})$. Notons que cela suppose connue une décomposition de G en produit direct de sous-groupes cycliques.

(2) Pour chaque morphisme ω, on calcule la matrice P_ω. Une famille de colonnes de P_ω linéairement indépendantes fournit une base de V_ω (notons que si $P_\omega = 0$, on obtient la famille vide).

(3) On recolle les bases obtenues pour obtenir la base voulue.

4.4. Remarque. Le lecteur aguerri aura remarqué que le corollaire 4.2 n'est rien d'autre que le fait que les caractères irréductibles forment une famille orthonormée pour un produit scalaire adéquat sur l'espace des fonctions centrales, et que, par ailleurs, les sous-espaces V_ω sont les composantes isotypiques liées à la représentation $\rho : G \longrightarrow \mathrm{GL}_n(F)$.

Il reste maintenant à comprendre comment calculer une $\mathbb{Z}$-base du groupe abélien Γ. On va pour cela s'appuyer sur l'algorithme de réduction des matrices à coefficients dans un anneau euclidien (théorème XX-2.3).

4.5. Proposition. *Soit $\mathscr{G} = \{g'_1, \ldots, g'_m\}$ une partie génératrice*[7] *de G et soit ζ_e une racine primitive e-ième de l'unité.*

Pour tout entier $i \in [\![1, m]\!]$, et pour tout entier $j \in [\![1, n]\!]$, soit $a_{ij} \in \mathbb{Z}$ tel que $\omega_j(g'_i) = \zeta_e^{a_{ij}}$. Soit $C = (a_{ij}) \in \mathrm{M}_{m \times n}(\mathbb{Z})$, et écrivons

$$UCV = \begin{pmatrix} \delta_1 & & & & & \\ & \ddots & & & & \\ & & \delta_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix},$$

avec $U \in \mathrm{GL}_m(\mathbb{Z})$, $V \in \mathrm{GL}_n(\mathbb{Z})$, $\delta_1, \ldots, \delta_r \in \mathbb{Z} \setminus \{0\}$.

7. On peut simplement prendre $\mathscr{G} = G$, par exemple.

Posons enfin $e_k = \dfrac{e}{\operatorname{pgcd}(\delta_k, e)}$ *pour tout* $k \in [\![1, r]\!]$, *et soient* $C_1, \ldots, C_n$ *les colonnes de* V. *Alors, les vecteurs*

$$\boldsymbol{\gamma}^{(1)} = e_1 C_1, \ \ldots, \ \boldsymbol{\gamma}^{(r)} = e_r C_r, \ \boldsymbol{\gamma}^{(r+1)} = C_{r+1}, \ldots, \ \boldsymbol{\gamma}^{(n)} = C_n$$

forment une $\mathbb{Z}$*-base de* Γ.

Démonstration. Rappelons que

$$\Gamma = \Big\{ \boldsymbol{\gamma} \in \mathbb{Z}^n \mid \omega_1(g)^{\gamma_1} \cdots \omega_n(g)^{\gamma_n} = 1, \ \text{pour tout } g \in G \Big\}.$$

Puisque $\omega_1, \ldots, \omega_n$ sont des morphismes de groupes et que $\mathscr{G}$ est une partie génératrice de G, il est facile de voir que

$$\Gamma = \Big\{ \boldsymbol{\gamma} \in \mathbb{Z}^n \ \Big| \ \omega_1(g_i')^{\gamma_1} \cdots \omega_n(g_i')^{\gamma_n} = 1, \ \text{pour tout } i \in [\![1, m]\!] \Big\}.$$

Comme ζ_e est d'ordre e, en gardant les notations de l'énoncé, on a ainsi

$$\Gamma = \Big\{ \boldsymbol{\gamma} \in \mathbb{Z}^n \ \Big| \ \sum_{j=1}^{n} a_{ij}\gamma_j \equiv 0 \ [e], \ \text{pour tout } i \in [\![1, m]\!] \Big\},$$

ce qui se récrit

$$\Gamma = \Big\{ \boldsymbol{\gamma} \in \mathbb{Z}^n \mid C\boldsymbol{\gamma} \equiv 0 \ [e] \Big\}.$$

Si $\boldsymbol{\gamma} \in \mathbb{Z}^n$, on pose $\boldsymbol{\gamma}' = V^{-1}\boldsymbol{\gamma}$. On a alors

$$C\boldsymbol{\gamma} \equiv 0 \ [e] \iff UC\boldsymbol{\gamma} \equiv 0 \ [e] \iff UCV\boldsymbol{\gamma}' \equiv 0 \ [e].$$

Si $\boldsymbol{\gamma}' = \begin{pmatrix} \gamma_1' \\ \vdots \\ \gamma_n' \end{pmatrix}$, cette dernière condition équivaut à

$$\delta_k \gamma_k' \equiv 0 \ [e], \quad \text{pour tout } k \in [\![1, r]\!].$$

Soit $\Delta_k = \operatorname{pgcd}(\delta_k, e)$. On peut alors écrire $\delta_k = \Delta_k u_k$ et $e = \Delta_k e_k$, où u_k et e_k sont premiers entre eux. En particulier, u_k est inversible modulo e_k. On a alors

$$\delta_k \gamma_k' \equiv 0 \ [e] \iff u_k \gamma_k' \equiv 0 \ [e_k] \iff \gamma_k' \equiv 0 \ [e_k],$$

la dernière équivalence découlant du fait que u_k est inversible modulo e_k.

Par conséquent, on obtient

$$\delta_k \gamma_k' \equiv 0 \ [e], \ \text{pour tout } k \in [\![1, r]\!] \iff \boldsymbol{\gamma}' = \begin{pmatrix} c_1 e_1 \\ \vdots \\ c_r e_r \\ c_{r+1} \\ \vdots \\ c_n \end{pmatrix}, \ c_1, \ldots, c_n \in \mathbb{Z}.$$

Cette dernière condition se réinterprète en disant que $\boldsymbol{\gamma}'$ est une combinai-

son $\mathbb{Z}$-linéaire de $e_1\varepsilon_1, \ldots, e_r\varepsilon_r, \varepsilon_{r+1}, \ldots, \varepsilon_n$. Il est aisé de voir que ces n vecteurs sont $\mathbb{Z}$-linéairement indépendants (ils le sont déjà sur $\mathbb{Q}$!).

Par conséquent, $\gamma \in \Gamma$ si, et seulement si, $\gamma' = V^{-1}\gamma$ s'écrit de manière unique comme combinaison $\mathbb{Z}$-linéaire de $e_1\varepsilon_1, \ldots, e_r\varepsilon_r, \varepsilon_{r+1}, \ldots, \varepsilon_n$. Comme $V \in \mathrm{GL}_n(\mathbb{Z})$, on en déduit que $\gamma \in \Gamma$ si, et seulement si, γ s'écrit de manière unique comme combinaison $\mathbb{Z}$-linéaire des vecteurs

$$e_1 V\varepsilon_1, \ \ldots, \ e_r V\varepsilon_r, \ V\varepsilon_{r+1}, \ \ldots, \ V\varepsilon_n,$$

ce qu'il fallait démontrer. □

Nous allons maintenant donner un algorithme qui, à partir d'un morphisme de groupes donné $\rho : G \longrightarrow \mathrm{GL}_n(F)$, permet de calculer des fractions rationnelles $f_1, \ldots, f_n$ G-invariantes algébriquement indépendantes sur F telles que $F(X_1, \ldots, X_n)^G = F(f_1, \ldots, f_n)$.

Algorithme

Étape 1. On calcule tout d'abord une matrice $R \in \mathrm{GL}_n(F)$ vérifiant

$$R^{-1}\rho(g)R = \begin{pmatrix} \omega_1(g) & & \\ & \ddots & \\ & & \omega_n(g) \end{pmatrix}, \quad \text{pour tout } g \in G,$$

où $\omega_1, \ldots, \omega_n : G \longrightarrow \mu_e(\overline{F})$ sont des morphismes de groupes (pas nécessairement distincts).

Pour cela, on peut procéder comme suit.

(i) On calcule tous les morphismes de groupes $G \longrightarrow \mu_e(\overline{F})$.
Si $G = \langle g_1\rangle \odot \cdots \odot \langle g_r\rangle$, avec $o(g_i) = d_i$, ce sont les morphismes $\omega_{z_1,\ldots,z_r}$, avec $z_i \in \mu_{d_i}(\overline{F})$, définis par

$$\omega_{z_1,\ldots,z_r}(g_1^{m_1} \cdots g_r^{m_r}) = z_1^{m_1} \cdots z_r^{m_r}.$$

(ii) Pour chaque morphisme $\omega : G \longrightarrow \mu_e(\overline{F})$, on calcule la matrice

$$P_\omega = \frac{1}{|G|} \sum_{g \in G} \omega(g)^{-1}\rho(g),$$

et l'on extrait un système de colonnes de P_ω linéairement indépendantes. On recolle le tout pour obtenir une matrice $R \in \mathrm{GL}_n(F)$, qui vérifie alors les propriétés souhaitées.

Étape 2. On pose

$$\Gamma = \left\{\gamma \in \mathbb{Z}^n \mid \omega_1(g)^{\gamma_1} \cdots \omega_n(g)^{\gamma_n} = 1, \ \text{pour tout } g \in G\right\},$$

où les ω_i sont les morphismes définis dans l'étape 1, et l'on calcule une base de Γ.

Pour ce faire, on peut effectuer les étapes suivantes.

(i) On choisit une partie génératrice $\mathscr{G} = \{g'_1, \ldots, g'_m\}$ de G (par exemple, on peut prendre les générateurs utilisés dans l'étape 1 (i), ou bien on peut prendre G lui-même).
Pour tout entier $i \in [\![1, m]\!]$, et pour tout entier $j \in [\![1, n]\!]$, soit $a_{ij} \in \mathbb{Z}$ tel que $\omega_j(g'_i) = \zeta_e^{a_{ij}}$, où ζ_e est une racine primitive e-ième de l'unité. On définit alors $C = (a_{ij}) \in \mathrm{M}_{m\times n}(\mathbb{Z})$.

(ii) En utilisant des opérations sur les lignes et les colonnes de C, on explicite alors $U \in \mathrm{GL}_m(\mathbb{Z})$, $V \in \mathrm{GL}_n(\mathbb{Z})$, $\delta_1, \ldots, \delta_r \in \mathbb{Z} \setminus \{0\}$ tels que

$$UCV = \begin{pmatrix} \delta_1 & & & & & \\ & \ddots & & & & \\ & & \delta_r & & & \\ & & & 0 & & \\ & & & & \ddots & \\ & & & & & 0 \end{pmatrix}.$$

Notons que l'on n'a pas besoin de calculer U ; seule la matrice V est importante. Celle-ci se calcule en faisant le produit de matrices de transvection et de permutation correspondant aux opérations sur les colonnes. On a donc intérêt à faire au maximum des opérations sur les lignes. On n'a pas besoin non plus d'avoir $\delta_1 \mid \cdots \mid \delta_r$, ce qui peut raccourcir un peu la mise en pratique de l'algorithme de réduction.

Posons $e_k = \dfrac{e}{\mathrm{pgcd}(\delta_k, e)}$, pour tout $k \in [\![1, r]\!]$. Si $C_1, \ldots, C_n$ sont les colonnes de V, alors les vecteurs

$$\gamma^{(1)} = e_1 C_1, \ \ldots, \ \gamma^{(r)} = e_r C_r, \ \ \gamma^{(r+1)} = C_{r+1}, \ \ldots, \ \gamma^{(n)} = C_n$$

forment une $\mathbb{Z}$-base de Γ.

Étape 3. On achève le calcul de $f_1, \ldots, f_n$.

(i) On considère la matrice R obtenue dans l'étape 1, et l'on calcule son inverse.

(ii) Soit $(\gamma^{(1)}, \ldots, \gamma^{(n)})$ une $\mathbb{Z}$-base de Γ. Si $R^{-1} = (a'_{ij})$, posons

$$U_j = \sum_k a'_{jk} X_k, \qquad \text{pour tout } j \in [\![1, n]\!].$$

Alors, les fractions rationnelles $f_j = U^{\gamma^{(j)}}$, $j \in [\![1, n]\!]$, sont algébriquement indépendantes sur F, et $F(X_1, \ldots, X_n)^G = F(f_1, \ldots, f_n)$.

4.6. Exemple. Reprenons le cas du groupe de Klein, à savoir

$$G = \langle s, t \rangle \subset \mathfrak{S}_4,$$

où $s = (1\ 2)(3\ 4)$ et $t = (1\ 3)(2\ 4)$, agissant sur $F[X_1, \ldots, X_4]$ par

$$\sigma \cdot P = P(X_{\sigma(1)}, \ldots, X_{\sigma(4)}).$$

D'après l'exemple 2.10 (1), cette action est celle induite par la représentation $\rho : G \longrightarrow \mathrm{M}_4(F)$ définie par

$$\rho(\sigma) = M_\sigma, \quad \text{pour tout } \sigma \in G.$$

On a donc

$$\rho(s) = M_s = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \quad \rho(t) = M_t = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix},$$

ainsi que

$$\rho(st) = M_{st} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

Comme $G = \langle s \rangle \odot \langle t \rangle$, et G est d'exposant 2, les morphismes $G \longrightarrow \mu_2(\overline{F})$ sont les quatre morphismes $\omega_1, \ldots, \omega_4$ caractérisés par

$$\begin{array}{llll} \omega_1(s) = 1 & \omega_2(s) = -1 & \omega_3(s) = 1 & \omega_4(s) = -1 \\ \omega_1(t) = 1 & \omega_2(t) = 1 & \omega_3(t) = -1 & \omega_4(t) = -1. \end{array}$$

On a donc

$$P_{\omega_1} = \frac{1}{4}(\mathrm{I}_4 + M_s + M_t + M_{st}) = \frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix},$$

$$P_{\omega_2} = \frac{1}{4}(\mathrm{I}_4 - M_s + M_t - M_{st}) = \frac{1}{4}\begin{pmatrix} 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \\ 1 & -1 & 1 & -1 \\ -1 & 1 & -1 & 1 \end{pmatrix},$$

$$P_{\omega_3} = \frac{1}{4}(\mathrm{I}_4 + M_s - M_t - M_{st}) = \frac{1}{4}\begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & 1 & -1 & -1 \\ -1 & -1 & 1 & 1 \\ -1 & -1 & 1 & 1 \end{pmatrix},$$

$$P_{\omega_4} = \frac{1}{4}(\mathrm{I}_4 - M_s - M_t + M_{st}) = \frac{1}{4}\begin{pmatrix} 1 & -1 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

On peut donc prendre

$$R = \frac{1}{4}\begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}.$$

Puisque $R^{-1} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \\ 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \end{pmatrix}$, on a ainsi

$$U_1 = X_1 + X_2 + X_3 + X_4, \quad U_2 = X_1 - X_2 + X_3 - X_4,$$
$$U_3 = X_1 + X_2 - X_3 - X_4, \quad U_4 = X_1 - X_2 - X_3 + X_4.$$

Ce calcul explique donc le choix des U_i dans l'exemple 2.2. En regardant l'action de G sur les U_i, on constate que

$$\sigma \cdot U_i = \omega_i(\sigma)^{-1} U_i, \quad \text{pour tout } \sigma \in G,$$

où les ω_i sont définis plus haut. Par conséquent, on a

$$\Gamma = \left\{ \gamma \in \mathbb{Z}^4 \mid \omega_1(\sigma)^{\gamma_1} \cdots \omega_4(\sigma)^{\gamma_4} = 1, \ \text{ pour tout } \sigma \in G \right\}.$$

Or, s et t sont des générateurs de G, et l'on a

$$\omega_1(s)^{\gamma_1} \cdots \omega_4(s)^{\gamma_4} = (-1)^{\gamma_2+\gamma_4}$$
$$\omega_1(t)^{\gamma_1} \cdots \omega_4(t)^{\gamma_4} = (-1)^{\gamma_3+\gamma_4}.$$

On en déduit facilement que $\gamma \in \Gamma$ si, et seulement si, γ est de la forme

$$\gamma = \begin{pmatrix} \gamma_1 \\ \gamma_4 + 2k \\ \gamma_4 + 2\ell \\ \gamma_4 \end{pmatrix}, \quad k, \ell, \gamma_1, \gamma_4 \in \mathbb{Z}.$$

La famille de vecteurs

$$\gamma^{(1)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix}, \quad \gamma^{(2)} = \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \end{pmatrix}, \quad \gamma^{(3)} = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 0 \end{pmatrix}, \quad \gamma^{(4)} = \begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

est donc une famille $\mathbb{Z}$-génératrice du groupe abélien Γ. Il est facile de voir qu'elle est aussi $\mathbb{Z}$-libre. C'est donc une $\mathbb{Z}$-base de Γ.

Finalement, on obtient que les fractions rationnelles U_1, U_2^2, U_3^2, $U_2U_3U_4$ sont algébriquement indépendantes, et que

$$F(X_1, X_2, X_3, X_4)^G = F(U_1, U_2^2, U_3^2, U_2U_3U_4),$$

comme annoncé précédemment.

Pour finir, nous allons maintenant donner un exemple où l'on met en œuvre la méthode proposée pour calculer une base de Γ.

4.7. Exemple

Soit F un corps vérifiant $\mathrm{car}(F) \neq 3$ et contenant une racine 9-ième de l'unité, que l'on notera ζ_9. On pose

$$s = \begin{pmatrix} \zeta_9^3 & 0 \\ 0 & \zeta_9^3 \end{pmatrix}, \quad t = \begin{pmatrix} \zeta_9^5 & 0 \\ 0 & \zeta_9^3 \end{pmatrix}.$$

Soit $G = \langle s, t \rangle$. On vérifie que s est d'ordre 3, t est d'ordre 9 et que

$$G = \langle s \rangle \odot \langle t \rangle \simeq \mathbb{Z}/3\mathbb{Z} \times \mathbb{Z}/9\mathbb{Z}.$$

En particulier, G est d'exposant 9. Enfin, on considère $\rho : G \longrightarrow \mathrm{GL}_2(F)$ le morphisme d'inclusion, ainsi que l'action de G associée sur $F(X_1, X_2)$.

Soient $\omega_1, \omega_2 : G \longrightarrow \mu_9(\overline{F})$ les morphismes, définis de manière unique, par

$$\begin{aligned} \omega_1(s) &= \zeta_9^3 & \omega_2(s) &= \zeta_9^3 \\ \omega_1(t) &= \zeta_9^5 & \omega_2(t) &= \zeta_9^3. \end{aligned}$$

Par définition de l'action associée à ρ, on a

$$g \cdot X_i \ = \ \omega_i(g)^{-1} X_i.$$

Posons $C = \begin{pmatrix} 3 & 3 \\ 5 & 3 \end{pmatrix} \in \mathrm{M}_2(\mathbb{Z})$. Nous allons maintenant utiliser des opérations élémentaires sur les lignes et les colonnes pour transformer C en une matrice diagonale. Nous calculerons la matrice V correspondant aux opérations sur les colonnes de proche en proche, en initialisant à $V = \mathrm{I}_2$, puis en multipliant à droite par la matrice correspondant à chaque opération élémentaire successive.

Si $M_1, M_2 \in \mathrm{M}_2(\mathbb{Z})$, on notera $M_1 \sim M_2$ lorsque l'on peut passer de la matrice M_1 à la matrice M_2 par une succession d'opérations élémentaires.

On commence par effectuer l'opération $L_2 \leftarrow L_2 - L_1$. On a donc

$$C \sim \begin{pmatrix} 3 & 3 \\ 2 & 0 \end{pmatrix}.$$

On échange ensuite les colonnes, si bien que

$$C \sim \begin{pmatrix} 3 & 3 \\ 0 & 2 \end{pmatrix},$$

et l'on fait $V \leftarrow V \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. Enfin, on effectue $C_2 \leftarrow C_2 - 3C_1$. Ainsi,

$$C \sim \begin{pmatrix} 3 & 0 \\ 0 & 2 \end{pmatrix},$$

et l'on fait $V \leftarrow V \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$.

Finalement, on a

$$V = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & -1 \end{pmatrix}.$$

Avec les notations de l'algorithme, on a donc

$$e_1 = \frac{9}{\operatorname{pgcd}(3, e)} = 3, \quad e_2 = \frac{9}{\operatorname{pgcd}(2, e)} = 9,$$

et les vecteurs

$$\gamma^{(1)} = 3 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 3 \end{pmatrix}, \quad \gamma^{(2)} = 9 \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 9 \\ -9 \end{pmatrix}$$

forment une $\mathbb{Z}$-base de Γ. Ainsi, X_2^3 et $X_1^9 X_2^{-9}$ sont algébriquement indépendantes sur F, d'où l'égalité $F(X_1, X_2)^G = F(X_2^3, X_1^9 X_2^{-9})$.

Bibliographie

[1] J.-M. Arnaudiès, J. Bertin, *Groupes, algèbres et géométrie, Tome 1.* Ellipses, 1993.

[2] J.-M. Arnaudiès, J. Bertin, *Groupes, algèbres et géométrie, Tome 2.* Ellipses, 1995.

[3] J.-M. Arnaudiès, H. Fraysse, *Cours de mathématiques, tome 1 : algèbre.* Dunod Université, 1987.

[4] Y. Amice, *Les nombres p-adiques.* PUF, 1975

[5] E. Artin, *Algèbre géométrique.* Éditions Jacques Gabay, 1962.

[6] M. Artin, *Algebra. Second edition.* Pearson, 2010.

[7] V. Beck, J. Malick, G. Peyré, *Objectif Agrégation. Deuxième édition.* Éditions H & K, 2005.

[8] J.-C. Carrega, *Théorie des corps : La règle et le compas. Nouvelle édition.* Coll. Formation des enseignants, Hermann, 2001.

[9] G. Berhuy, *Modules : théorie, pratique... et un peu d'arithmétique.* Mathématiques en devenir **109**, Calvage & Mounet, 2012.

[10] G. Berhuy, *Jeux d'ampoules (ou comment éviter la crise de nerfs à un électricien dépressif à coup d'algèbre linéaire sur* $\mathbb{F}_2$*).* Quadrature **79**, 32–42, Janvier-Février-Mars 2011.

[11] G. Berhuy, *Corps gauches et codes Wifi.* Quadrature **85**, 33–41, Janvier-Février-Mars 2015.

[12] G. Berhuy, F. Oggier, *An introduction to central simple algebras and their applications to wireless communication.* Mathematical Surveys and Monographs **191**, AMS (2013)

[13] E.A. Bender, *Characteristic polynomials of symmetric matrices.* Pacific J. Math. **25** (1968), 433-431

[14] E.A. Bender, *The dimension of symmetric matrices with a given minimal polynomial.* Linear Alg. and Appl. **3** (1970), 115–123

[15] N. Bourbaki, *Algèbre, chapitres 1 à 3.* Springer, 2007.

[16] Thomas Chomette, *Les colliers de G. Polyà.* Disponible sur http ://culturemath.ens.fr/maths/pdf/combi/polya.pdf

[17] Thomas Chomette, *Construction des polygones réguliers.* Disponible sur http ://culturemath.ens.fr/maths/pdf/geometrie/polygones.pdf

[18] A. Chambert-Loir, *Algèbre corporelle.* Éditions de l'École Polytechnique, 2005.

[19] H. Cohen, *A Course in Computational Number Theory.* Graduate Texts in Mathematics **138**, Springer, 1993.

[20] P. Colmez, *Éléments d'analyse et d'algèbre (et de théorie des nombres). Nouvelle édition.* Éditions de l'École Polytechnique, Ellipses, 2012.

[21] Keith Conrad, *Transitive Group Actions.* Disponible à l'adresse : https ://math.uconn.edu/ kconrad/blurbs/grouptheory/transitive.pdf

[22] M. Coste, *Algèbre linéaire sur les entiers.* Disponible sur http ://agreg-maths.univ-rennes1.fr/documentation/docs/alglinent.pdf

[23] M. Coste, *Jeu de taquin et générateurs du groupe alterné.* Disponible à l'adresse :
http ://agreg-maths.univ-rennes1.fr/documentation/docs/taquin.pdf

[24] A. Debreil, *Groupes finis et treillis de leurs sous-groupes.* Mathématiques en devenir **114**, (deuxième édition). Calvage & Mounet, 2025.

[25] E. Driver, P.A. Leonard, K. S. Williams, *Irreducible Quartic Polynomials with Factorizations Modulo p.* American Mathematical Monthly, Vol. **112**, No. 10 (2005), 876–890

[26] C. de Séguins Pazzis, *Invitation aux formes quadratiques.* Mathématiques en devenir **104**, Calvage et Mounet, 2010

[27] J.-P. Escofier, *Galois theory.* Graduate Texts in Mathematics **204**, Springer, 2000.

[28] D. E. Fields, *Zero divisors and nilpotent elements in power series rings.* Proceedings of the AMS Vol. 27 No 3 (1971), 427–433.

[29] F. Gouvêa, *p-adic numbers : an introduction. Second edition.* Universitext, Springer, 1997.

[30] G. Gras, M.-N. Gras, *Algèbre fondamentale, arithmétique.* Mathématiques à l'université. Cours et exercices corrigés. Ellipses, 2004.

[31] J. Grifone, *Algèbre linéaire.* Cepadues éditions, 1990.

[32] F. Krakowski, *Eigenwerte und Minimalpolynome symmetrischer Matrizen in kommutativen Körpen.* Comm. Math. Helv. **32** (1957), 224–240

[33] J.-L. Krivine, *Théorie des ensembles.* Cassini, 2007.

[34] S. Lang, *Algebra. Revised third edition.* Graduate Texts in Mathematics, Springer, 2005.

[35] K. Lux, H. Pahlings, *Representations of groups : a computational approach.* Cambridge studies in advanced mathematics **124**, Cambridge university press, 2010.

[36] P. Ortiz, *Exercices d'algèbre.* Maths Agreg. Ellipses, 2004.

[37] D. Perrin, *Cours d'algèbre.* Maths Agreg. Ellipses, 1996.

[38] G. Rauch, *Les groupes finis et leurs représentations.* Mathématiques 2e cycle. Cours et exercices corrigés. Ellipses, 2000.

[39] J. J. Rotman, *An introduction to the theory of groups. Fourth edition.* Graduate Texts in Mathematics **148**, Springer, 1994.

[40] G. Rousseau, *On the quadratic reciprocity law.* J. Austral. Math. Soc. (Series A) **51** (1991), 423–425.

[41] P. Samuel, *On unique factorization domains.* Illinois J. Math. Vol. **5** (1961), 1–17.

[42] L. Schwartz, *Analyse I.* Coll. Enseignement des sciences, **42**. Hermann, Paris, 1995.

[43] J.-P. Serre, *Cours d'arithmétique (quatrième édition).* PUF, 1994.

[44] J.-P. Serre, *Groupes finis – Cours au Collège de France.* Disponible sur https ://www.college-de-france.fr/media/jean-pierre-serre/UPL2937151343298039815_1___Groupes_finis.pdf

[45] J.-P. Serre, *Linear representations of finite groups.* Graduate Texts in Mathematics **42**, Springer, 1977.

[46] I. Stewart, *Galois theory, Fourth edition.* CRC Press, 2015.

[47] P. Tauvel, *Corps commutatifs et théorie de Galois.* Mathématiques en devenir **102**, Calvage & Mounet, 2008.

[48] J.-P. Tignol, *Galois theory of algebraic equations.* World Scientific, 2001.

[49] F. Ulmer, *Théorie des groupes.* Ellipses, 2012.

Index

action
  2-transitive, 332
  d'un groupe sur un ensemble, 171
  fidèle, 172
  linéaire, 1161
  primitive, 322
affinité, 91
d'Alembert (théorème de), 848
algèbre
  d'un groupe, 1168
  de quaternions, 593
  de type fini, 709
  sur un corps, 39
algébriquement indépendants, 1269
algorithme
  d'Euclide, 25, 548
  de Berlekamp, 695
  de réduction des matrices, 609
anneau
  à division, 414
  à pgcd, 576
  commutatif, 14, 403
  de Bézout, 580
  de Boole, 460
  définition, 13, 403
  des polynômes, 433
  euclidien, 545
  factoriel, 558, 559
  intégralement clos, 867
  intègre, 414
  noethérien, 553
  principal, 423
  produit direct, 404
  quotient, 472
  trivial, 405
annulateur d'un endomorphisme, 978
application
  bijective, 6
  définition, 4
  injective, 5
  surjective, 6
application linéaire, 621
  définition, 38
  matrice représentative, 46
Artin (lemme d'), 948
automorphisme
  d'espaces vectoriels, 39
  d'un groupe, 131
  intérieur, 133

base, 41, 623
  d'un groupe, 266
  de Jordan, 1011
  directe,indirecte, 115
  duale, 1108
Bender (théorème de), 1140
Berlekamp (algorithme de), 695
Bézout
  anneau de, 580
  relation de, 146
  théorème de, 26
bicommutant, 1072
bloc de Jordan, 1010
Boole (anneau de), 460
Burnside, 1249, 1250
  formule de, 178
  théorème de, 1251

caractère
  d'une représentation, 1201
  linéaire, 354, 1162
caractéristique d'un anneau, 408
carré symétrique/alterné d'une représentation, 1184
Cauchy (théorème de), 181
Cayley (théorème de), 179
Cayley-Hamilton (théorème de), 982
cellule de Jordan, 1010
centralisateur, 392
centre
  d'un anneau, 406
  d'un groupe, 138

classe
de conjugaison, 133
classe d'équivalence, 7
classe modulo un sous-groupe, 147
clôture
algébrique, 857
galoisienne, 917
normale, 914
comatrice, 73
combinaison linéaire, 621
commutant, 1072
commutateur, 327
complété I-adique, 764
composante isotypique, 1239
compositum, 816
concaténation, 279
congruence modulo un entier, 30
congruentes (matrices), 1145
conjugaison, 133
contenu d'un polynôme, 658
corps, 414
algébriquement clos, 857
de décomposition, 853
de rupture, 850
définition, 14
des fractions, 418
des racines, 853
parfait, 893
correspondance, 3
critère
d'Iwasawa, 329
de Wantzel, 811, 821
critère d'irréductibilité
d'Eisenstein, 653
par réduction modulo un idéal, 651
crochet de dualité, 1085
cycle, 204

d'Alembert (théorème de), 848
développement de Hensel, 784
développement en série formelle, 749
décomposition
de Dunford, 1004
de Frobenius, 1055
degré
d'un élément, 815
d'une extension, 801
d'une représentation, 1162
partiel, 451
déterminant
d'un endomorphisme, 68
d'une matrice, 69, 603
de n vecteurs, 66
minimum d'un code , 589
dilatation, 79
dimension, 43
discriminant, 847
diviseur de zéro, 411
divisibilité dans un anneau, 531
Dunford (décomposition de), 1004

Eisenstein
critère d'irréductibilité, 653
polynôme d', 654
élément
algébrique, 812
entier, 864
exprimable par radicaux, 957
indivisible, 583
inversible (d'un anneau), 410
irréductible, 534
maximal, 15
nilpotent, 412
normalisé, 562
primitif, 888
régulier, 499, 778
séparable, 881
totalement réel, 1126
transcendant, 812
éléments
étrangers, 483
associés, 532
premiers entre eux, 532

endomorphisme
adjoint, 97, 1114
antiautoadjoint, 97
autoadjoint, 97
cyclique, 1050
d'un espace vectoriel, 38
diagonalisable, 990
nilpotent, 998
normal, 97
semi-simple, 1037
simple, 1037
transposé, 1116
trigonalisable, 995
unitaire, 97
ensemble
inductif, 16
quotient, 7
totalement ordonné, 15
entier b-adique, 765
espace vectoriel, 37
de type fini, 43
dual, 48
euclidien, 95
hermitien, 95
orienté, 115
Euclide
algorithme, 25, 548
lemme, 27, 541, 560
Euler (indicatrice), 158
évaluation d'un polynôme, 439
exposant, 351
extension de corps
2-décomposable, 810
abélienne, 931
algébrique, transcendante, 812
cyclique, 931
définition, 801
de Kummer, 935
galoisienne, 915
monogène, 804
normale, 912
radicale, 956
résoluble, 957
séparable, 881
extension des scalaires, 1185
extensions de corps linéairement disjointes, 819

famille
génératrice, 41, 621
libre,liée, 40, 622
Fermat
nombre de, 824
théorème, 181
Fisher (théorème de), 1280
fonction
centrale, 1202
monomiale, 1278
polynomiale, 1278
forme
alternée, 64
antisymétrique, 64
bilinéaire
définition, 1086
matrice représentative, 1104
multiplicative, 1148
non dégénérée, 1089
de Jordan, 1011
linéaire, 48
n-linéaire, 64
normale de Smith, 612
symétrique, 64
formule
d'inversion de Möbius, 155
de Burnside, 178
de réversion de Lagrange, 738
Frobenius
décomposition de, 1055
théorème de, 181, 1224
théorème de réduction, 1055

G-équivariant, 288
Galois
groupe d'un polynôme, 921
groupe d'une extension, 825
théorème, 951
Gauss
lemme de, 26, 540, 561
théorème de, 829

groupe
abélien, 127
abélien, 12
abélien libre, 358
alterné, 63, 217
cyclique, 156
d'exposant fini, 351
définition, 12, 127
dérivé, 327
divisible, 372
libre, 267
linéaire, 39
monogène, 156
orthogonal, 97
p-primaire, 348
produit direct externe, 128, 299
projectif spécial linéaire, 332
quotient, 239
résoluble, 261
simple, 142
spécial linéaire, 69
spécial orthogonal, 101
super-résoluble, 261
symétrique, 59, 128
unitaire, 97
groupe de Galois
d'un polynôme, 921
d'une extension, 825

Hensel
lemme de, 776
Hilbert
théorème, 706
théorème 90, 936
hyperplan, 50

idéal
à droite, 420
à gauche, 420
bilatère, 420
engendré par une partie, 422
maximal, 475
premier, 475
principal, 423
idempotent, 412
central, 427
système fondamental d'idempotents orthogonaux, 427
image, 3, 6
directe, 6
réciproque, 6
indépendants
algébriquement, 1269
linéairement, 40
indicatrice d'Euler, 158
indice d'un sous-groupe, 148
invariants de similitude, 1055
inverse d'un élément, 127
isométrie, 97
directe, 101
indirecte, 101
isomorphisme
d'anneaux, 407
d'espaces vectoriels, 39
de groupes, 131
Iwasawa
critère de, 329
théorème de, 326

Jacobi (symbole de), 522
Jordan
bloc de, 1010
cellule de, 1010
forme de, 1011
théorème de réduction, 1015
Jordan-Hölder
suite de, 253
théorème de, 260

Kummer
extension de corps, 935
théorème de, 936

Lagrange
formule de réversion de, 738
Legendre (symbole de), 515

lemme
d'Artin, 948
d'Euclide, 27, 541, 560
d'indépendance de Dedekind, 878
de décomposition des noyaux, 977
de Gauss, 26, 540, 561
de Hensel, 776
de Schur, 1160, 1210
de Zorn, 16
du papillon, 256

Maschke (théorème de), 1188
matrice
anticirculante, 87
antihermitienne, 98
antisymétrique, 98
circulante, 87
compagnon, 1051
de dilatation, 81, 608
de passage, 47
de transposition, 608
de transposition signée, 616
de transvection, 78, 608
de Vandermonde, 86
hermitienne, 98
image, 629
noyau, 629
orthogonale, 98
représentative, 46
symétrique, 98
unitaire, 98
matrices congruentes, 1145
maximum, 15
mineur, 71
Möbius (formule d'inversion de), 155
module
de type fini, 621
définition, 618
indécomposable, 1159
libre, 623
noethérien, 640
simple, 1160
monoïde, 878
monôme, 450

morphisme
d'algèbres, 39
d'anneaux, 407
d'évaluation, 272, 1269
d'extensions, 801
de groupes, 130
de monoïdes, 878
de représentations, 1170
mot
définition, 279
longueur, 279

nombre de Fermat, 824
normalisateur, 391
norme, 895
noyau
d'un caractère, 1209
d'un morphisme de groupes, 137

orbite sous l'action
d'un élément, 177
d'un groupe, 175
d'une permutation, 204
ordre
d'un élément, 151
d'un groupe, 130
d'une série formelle, 716
orthogonal, 95, 1087

p-groupe, 182
p-Sylow, 317
Pólya (théorème de), 189
partie génératrice, 144
partition
d'un ensemble, 8
d'un entier, 1079
permutation, 59, 201
pgcd, 533
plongement, 878
point fixe, 175

polynôme
d'Eisenstein, 654
annulateur, 979
caractéristique
d'un endomorphisme, 982
d'une matrice, 980
cyclotomique, 675
d'endomorphisme, 976
degré, 433
dérivé, 441
homogène, 450
minimal
d'un élément algébrique, 813
d'un endomorphisme, 979
primitif, 650
résoluble par radicaux, 957
séparable, 881
symétrique
définition, 449
élémentaire, 450
unitaire, 433
ppcm, 533
présentation d'un groupe, 276
produit
mixte, 116
vectoriel, 118
produit direct
externe, 299
interne, 299, 303
produit scalaire, 95
produit semi-direct
externe, 306
interne, 306
produit tensoriel
d'espaces vectoriels, 1178
de représentations, 1183
projecteur, 94
spectral, 1001
projection
canonique, 9, 472
orthogonale, 97

racine
multiple, 442
ordre, 442
simple, 442
racine d'un polynôme, 439
rang d'un groupe abélien libre, 361
réflexion
définition, 92
orthogonale, 103
relation
d'équivalence, 7
d'équivalence
grossière, 322
d'ordre, 15
définition, 7
de Bézout, 26, 146
relation d'équivalence
G-stable, 321
renversement, 106
représentation linéaire
définition, 1162
de permutation, 1166
duale, 1183
fidèle, 1162
inflation, 1187
irréductible, 1173
matricielle, 1163
quotient, 1187
régulière, 1166
restriction, 1186
somme directe externe, 1182
somme directe interne, 1182
triviale, 1162

s.c.r.i, 544
Schreier (théorème de), 258
Schur
lemme de, 1160, 1210
théorème de, 736
série formelle
composition, 726
définition, 432
dérivée, 734
famille sommable, 723
réverse, 734
signature, 62, 216, 224
Smith (forme normale de), 612
somme directe, 44
interne, 1157
orthogonale, 95

somme orthogonale de formes bilinéaires, 1146
sous-anneau, 406
à division, 416
sous-corps, 416
sous-espace vectoriel
caractéristique, 1000
définition, 38
engendré par une partie, 38
propre, 990
sous-extension
définition, 801
engendrée par une partie, 804
sous-groupe, 13
maximal, 290, 325
transitif, 235
sous-module
définition, 619
engendré par une partie, 620
sous-mot, 280
sous-représentation, 1173
sous-groupe
définition, 135
distingué, 137
engendré par une partie, 143
propre, 135
strict, 135
trivial, 135
spectre d'un endomorphisme, 990
stabilisateur, 174
Steinitz (théorème de), 861
suite de composition
définition, 248
équivalence, 255
longueur, 248
quotients, 248
raffinement, 251
suite de Jordan-Hölder, 253
suite exacte
définition, 308
scindée, 308
supplémentaire, 45
support d'une permutation, 202
Sylow (théorème de), 319
symbole
de Jacobi, 522
de Legendre, 515
symétrie, 94
orthogonale, 97

table des caractères, 1230, 1232, 1234, 1236
taquin (jeu de), 228
théorème
90 de Hilbert, 936
chinois, 481, 483
d'existence de la signature, 215
d'isomorphisme (deuxième), 244
d'isomorphisme (premier), 243, 474, 1096
d'isomorphisme (second), 475
d'isomorphisme (troisième), 245
d'Iwasawa, 326
de Bézout, 26
de Bender, 1140
de Burnside, 1249, 1250
de Cauchy, 181
de Cayley, 179
de Cayley-Hamilton, 982
de complétude I-adique, 771
de d'Alembert, 848
de diagonalisation simultanée, 994
de diagonalisation (1), 991
de diagonalisation (2), 993
de division de polynômes généralisée, 434
de factorisation
anneaux, 473
espaces vectoriels, 1095
groupes, 241
de Fermat, 181
de Fisher, 1280
de Frobenius, 181, 1224
de Gauss, 661
de Gauss (polygones réguliers

constructibles), 829
de Hilbert, 706
de Hilbert (90), 936
de Jordan (réduction), 1015
de Jordan-Hölder, 260
de Krull, 476
de Kummer, 936
de l'élément primitif, 887
de la base adaptée, 627
de la base incomplète, 41
de la correspondance de Galois, 951
de la division euclidienne, 23
de Lagrange, 150
de Maschke, 1188
de Pólya, 189
de prolongement des morphismes d'anneaux, 850
de réciprocité quadratique, 517
de réduction
de Frobenius, 1055
de Jordan, 1015
des endomorphismes normaux, 110
des matrices, 609
de Schreier, 258
de Schur, 736
de Steinitz, 861
de structure
des corps finis, 690
des groupes abéliens de type fini, 371, 637
des groupes abéliens divisibles, 382
des polynômes symétriques, 451
de Sylow, 319
de Wedderburn, 683
du raffinement, 258
fondamental de l'arithmétique, 27
$p^a q^b$ de Burnside, 1251
torsion, 347
trace, 895
transposition, 59, 204
transvection, 75
type d'une permutation, 213

valeur propre
définition, 990
multiplicité, 991
valuation
π-adique, 559
p-adique, 28
Vandermonde (matrice de), 86
vecteur
cyclique, 1050
propre, 990
u-maximal, 1050
vecteurs orthogonaux, 95

Wantzel (critère de), 811, 821
Wedderburn (théorème de), 683

Zorn (lemme de), 16

Chez le même éditeur,
la collection *Nano*

101. — Benoît Kloeckner. Un bref aperçu de la géométrie projective
102. — Michel Balazard. *Le théorème des nombres premiers*
103. — Bruno Kahn. *Fonctions zêta et L de variétés et de motifs*
104. — Patrick Dehornoy. *Le calcul des tresses*
105. — Alain Debreil, Rached Mneimné. *Le groupe symétrique $\mathfrak{S}_4$ et ses métamorphoses*
106. — Roger Mansuy, *Introduction aux graphes aléatoires (et à la méthode probabiliste)*
107. — Jean-Denis *Eiden, Espaces euclidiens, avec une ouverture vers les espaces préhilbertiens réels*
108. — François Berteloot. *Les familles normales*
109. — Antoine Chambert-Loir. *Théorie de l'information.* Trois théorèmes de Claude Shannon
110. — Bernard Candelpergher. *Les séries, d'Euler à Ramanujan*
111. — François Rouvière. *Transformations de Radon. Cinq leçons de géométrie intégrale*
112. — Emmanuel Amiot. *Une introduction aux mathématiques de la musique*
113. — Thierry Coquand, Henri Lombardi. *Résolutions libres finies. Méthodes constructives*
114. — Claude Dellacherie. *Outils de théorie du potentiel*
115. — Anna Cadoret. *Modules sur un anneau et représentations linéaires des groupes finis*
116. — Alain Debreil, Rached Mneimné. *Les groupes de frise, avec un cours intensif de géométrie supérieure* (à paraître)

Imprimé dans l'Union européenne par les soins de PrintCorpgroup, Paris
Dépôt légal : juin 2025